



05-10-01

09831455 .050801

JC04 Rec'd PCT/PTO 08 MAY 2001

PTO-1390 U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371		ATTORNEY'S DOCKET NUMBER PF-0634 USN
		U.S. APPLICATION NO. (If known, see 37 CFR 1.5) TO BE ASSIGNED 09/831455
INTERNATIONAL APPLICATION NO. PCT/US99/27009	INTERNATIONAL FILING DATE 12 November 1999	PRIORITY DATE CLAIMED 12 November 1998
TITLE OF INVENTION HUMAN HYDROLASE PROTEINS		
APPLICANT(S) FOR DO/EO/US INCYTE PHARMACEUTICALS, INC.; TANG, Y. Tom; HILLMAN, Jennifer L.; YUE, Henry; LAL, Preeti; BANDMAN, Olga; CORLEY, Neil C.; GUEGLER, Karl J.; BAUGHN, Mariah R.; LU, Dyung Aina M.; AZIMZAI, Yalda; YANG, Junming		
<p>Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:</p> <ol style="list-style-type: none"><input checked="" type="checkbox"/> This is the FIRST submission of items concerning a filing under 35 U.S.C. 371.<input type="checkbox"/> This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371.<input type="checkbox"/> This is an express request to promptly begin national examination procedures (35 U.S.C. 371 (f)).<input type="checkbox"/> The US has been elected by the expiration of 19 months from the priority date (PCT Article 31).<input checked="" type="checkbox"/> A copy of the International Application as filed (35 U.S.C. 371(c)(2))<ol style="list-style-type: none"><input type="checkbox"/> is attached hereto (required only if not communicated by the International Bureau)<input type="checkbox"/> has been communicated by the International Bureau.<input checked="" type="checkbox"/> is not required, as the application was filed in the United States Receiving Office (RO/US).<input type="checkbox"/> An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)).<input checked="" type="checkbox"/> Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))<ol style="list-style-type: none"><input type="checkbox"/> are attached hereto (required only if not communicated by the International Bureau).<input type="checkbox"/> have been communicated by the International Bureau.<input type="checkbox"/> have not been made; however, the time limit for making such amendments has NOT expired.<input checked="" type="checkbox"/> have not been made and will not be made.<input type="checkbox"/> An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).<input checked="" type="checkbox"/> An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).<input type="checkbox"/> An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).		
Items 11 to 16 below concern document(s) or information included:		
<ol style="list-style-type: none"><input type="checkbox"/> An Information Disclosure Statement under 37 CFR 1.97 and 1.98.<input type="checkbox"/> An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.27 and 3.31 is included.<input type="checkbox"/> A FIRST preliminary amendment. <input type="checkbox"/> A SECOND or SUBSEQUENT preliminary amendment.<input type="checkbox"/> A substitute specification.<input type="checkbox"/> A change of power of attorney and/or address letter.<input checked="" type="checkbox"/> Other items or information:<ol style="list-style-type: none">Transmittal Letter (2 pp, in duplicate)Return PostcardExpress Mail Label No.: EL 856 113 138 USRequest to Transfer		

HUMAN HYDROLASE PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of hydrolase proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative, autoimmune/inflammatory, neurological, renal, adrenal, and genetic disorders.

BACKGROUND OF THE INVENTION

Hydrolysis is the breaking of a covalent bond in a substrate by introduction of a water molecule. The reaction involves a nucleophilic attack by the water molecule's oxygen atom on a target bond in the substrate. The water molecule is split across the target bond, breaking the bond and generating two product molecules. Hydrolases participate in reactions essential to functions such as cell signaling, cell proliferation, inflammation, apoptosis, secretion and excretion. Hydrolases are involved in key steps in disease processes involving these functions. Hydrolases, or hydrolytic enzymes, may be grouped by substrate specificity into classes including aminohydrolases, phospholipases, carboxyl-esterases, phosphodiesterases, glycosidases, glyoxalases, sulfatases, phosphohydrolases, serine hydrolases, and lysozymes.

NG,NG-dimethylarginine dimethylaminohydrolase (DDAH) is an enzyme that hydrolyzes the endogenous nitric oxide synthase (NOS) inhibitors, NG-monomethyl-arginine and NG,NG-dimethyl-L-arginine, to L-citrulline. Inhibiting DDAH can cause increased intracellular concentration of NOS inhibitors to levels sufficient to inhibit NOS. Therefore, DDAH inhibition may provide a method of NOS inhibition, and changes in the activity of DDAH could play a role in pathophysiological alterations in nitric oxide generation (MacAllister, R.J. et al. (1996) Br. J. Pharmacol. 119:1533-1540). DDAH was found in neurons displaying cytoskeletal abnormalities and oxidative stress in Alzheimer's disease. In age-matched control cases, DDAH was not found in neurons. This suggests that oxidative stress- and nitric oxide-mediated events play a role in the pathogenesis of Alzheimer's disease (Smith, M.A. et al. (1998) Free Rad. Biol. Med. 25:898-902).

Phosphodiesterases catalyze the hydrolysis of one of the two ester bonds in a phosphodiester compound. Phosphodiesterases are, therefore, crucial to a variety of cellular processes. Phosphodiesterases include DNA and RNA endo- and exo-nucleases, which are essential to cell growth and replication as well as protein synthesis.

Pancreatic lipase and colipase form a complex that plays a key role in dietary fat digestion by converting insoluble long chain triacylglycerols into more polar molecules able to cross the brush border of intestinal cells. Colipase binds to the C-terminal domain of lipase. In solution, this

interaction involves the formation of an ion pair between a glutamic acid residue of colipase and a lysine residue of lipase. These residues are strictly conserved among species (Ayvazian, L. et al. (1998) J. Biol. Chem. 273:33604-33609). Colipase appears to overcome the inhibitory effects of bile salts on pancreatic lipase (OMIM 246600 on April 28, 1999).

5 Carboxylesterases are proteins that hydrolyze carboxylic esters and are classified into three categories- A, B, and C. Most type-B carboxylesterases are evolutionarily related and are considered to comprise a family of proteins. The type-B carboxylesterase family of proteins includes vertebrate acetylcholinesterase, mammalian liver microsomal carboxylesterase, mammalian bile-salt-activated lipase, and duck fatty acyl-CoA hydrolase. Some members of this protein family are not catalytically
10 active but contain a domain related evolutionarily to other type-B carboxylesterases, such as thyroglobulin and the Drosophila protein neuractin. The active site of carboxylesterases involves three residues: a serine, a glutamate or aspartate, and a histidine. The sequence surrounding this catalytic site is well conserved and can be used as a signature pattern (PROSITE: PDOC00112 at www.expasy.ch/cgi-bin/get-prodoc-entry on May 11, 1999).

15 Acyl-CoA thioesterase is another member of the carboxylesterase family (Alexson, S.E. et al. (1993) Eur. J. Biochem. 214:719-727). Evidence suggests that acyl-CoA thioesterase has a regulatory role in steroidogenic tissues (Finkelstein, C. et al. (1998) Eur. J. Biochem. 256:60-66).

A phospholipase A₂ inhibitor has been identified that has 33% sequence homology with human leucine-rich α_2 -glycoprotein (Okumura, K. et al. (1998) J. Biol. Chem. 273:19469-19475).

20 Leucine-rich repeat (LRR) consensus sequences have also been found in the primary structure of many proteins, including proteins that participate in biologically important processes, such as receptors for hormones, enzymes, enzyme inhibitors, proteins for cell adhesion, and ribosome-binding proteins. All proteins containing LRR domains are thought to be involved in protein-protein interactions.

25 The glyoxylase system consists of glyoxalase I, which catalyzes the formation of S-D-lactoylglutathione from methyglyoxal, a side product of triose-phosphate energy metabolism, and glyoxylase II, which hydrolyzes S-D-lactoylglutathione to D-lactic acid and reduced glutathione. Methyglyoxal levels are elevated during hyperglycemia, likely due to increased triose-phosphate energy metabolism. Elevated levels of glyoxylase II activity have been found in human non-insulin-
30 dependent diabetes mellitus and in a rat model of this disease. The glyoxylase system has been implicated in the detoxification of bacterial toxins, and in the control of cell proliferation and microtubule assembly. Elevated levels of S-D-lactoylglutathione, the substrate of glyoxylase II, induced growth arrest and toxicity in HL60 cells. Thus, the glyoxylase system, and glyoxylase II in particular, may be associated with cell proliferation and autoimmune disorders such as diabetes.

The alpha/beta hydrolase protein fold is common to several hydrolases of diverse phylogenetic origin and catalytic function. Enzymes with the alpha/beta hydrolase fold have a common core structure consisting of eight beta-sheets connected by alpha-helices. The most conserved structural feature of this fold is the loops of the nucleophile-histidine-acid catalytic triad.

5 The histidine in the catalytic triad is completely conserved, while the nucleophile and acid loops accommodate more than one type of amino acid (Ollis, D.L. et al. (1992) Protein Eng. 5:197-211).

Sulfatases are members of a highly conserved gene family that share extensive sequence homology and a high degree of structural similarity. Sulfatases catalyze the cleavage of sulfate esters. To perform this function, sulfatases undergo a unique post-translational modification in the
10 endoplasmic reticulum that involves the oxidation of a conserved cysteine residue. A human disorder called multiple sulfatase deficiency is due to a defect in this post-translational modification step, leading to inactive sulfatases (Recksick, M. et al. (1998) J. Biol. Chem. 273:6096-6103).

Phosphohydrolases are enzymes that hydrolyze phosphate esters. Some phosphohydrolases contain a mutT domain signature sequence. MutT is a protein involved in the GO system responsible
15 for removing an oxidatively damaged form of guanine from DNA. A region of about 40 amino acid residues, found in the N-terminus of mutT, is also found in other proteins, including some phosphohydrolases (PROSITE: PDOC00695 at www.expasy.ch/cgi-bin/get-prodoc-entry on April 27, 1999).

Glycosidases catalyze the cleavage of hemiacetyl bonds of glycosides, which are compounds
20 that contain one or more sugar. Mammalian beta-galactosidase removes the terminal galactose from gangliosides, glycoproteins, and glycosaminoglycans. Beta-galactosidases belong to family 35 in the classification of glycosyl hydrolases. Deficiency of this enzyme is associated with the genetic disease GM1-gangliosidosis, also known as Morquio disease type B (PROSITE: PDOC00910 at www.expasy.ch/cgi-bin/get-prodoc-entry on May 12, 1999).

25 Serine hydrolases are a functional class of hydrolytic enzymes that contain a serine residue in their active site. This class of enzymes contains proteinases, esterases, and lipases which hydrolyze a variety of substrates and, therefore, have different biological roles. Proteins in this superfamily can be further grouped into subfamilies based on substrate specificity or amino acid similarities (Puente, X.S. and C. Lopez-Ont (1995) J. Biol. Chem. 270:12926-12932). One member of the serine hydrolase
30 superfamily is kraken, a Drosophila gene isolated from a Drosophila embryo cDNA library. Kraken belongs to a subfamily whose members catalyze cleavage of substrates with a carbonyl-containing group (Chan, E. et al. (1998) Gene 222:195-201).

The lysozyme c superfamily consists of conventional lysozymes c, calcium-binding lysozymes c, and α -lactalbumin (Prager, E.M. and P. Jolles (1996) EXS 75:9-31). The proteins in this

superfamily have 35-40% sequence homology and share a common three-dimensional fold, but can have different functions. Lysozymes c are ubiquitous in a variety of tissues and secretions and can lyse the cell walls of certain bacteria (McKenzie, H.A. (1996) EXS 75:365-409). Alpha-lactalbumin is a metallo-protein that binds calcium and participates in the synthesis of lactose (Iyer, L.K. and P.K. Qasba (1999) Protein Eng. 12:129-139). Alpha-lactalbumin occurs in mammalian milk and colostrum (McKenzie, supra).

Lysozymes catalyze the hydrolysis of certain mucopolysaccharides of bacterial cell walls, specifically, the beta (1-4) glycosidic linkages between N-acetylmuramic acid and N-acetylglucosamine, and cause bacterial lysis. Lysozymes occur in diverse organisms including viruses, birds, and mammals. In humans, lysozymes are found in spleen, lung, kidney, white blood cells, plasma, saliva, milk, tears, and cartilage (Online Mendelian Inheritance in Man (OMIM) #153450 Lysozyme; Weaver, L.H. et al. (1985) J. Mol. Biol. 184:739-741). Lysozyme c functions in ruminants as a digestive enzyme, releasing proteins from ingested bacterial cells, and may perform the same function in human newborns (Braun, O.H. et al. (1995) Klin. Pediatr. 207:4-7).

The two known forms of lysozymes, chicken-type and goose-type, were originally isolated from chicken and goose egg white, respectively. Chicken-type and goose-type lysozymes have similar three-dimensional structures, but different amino acid sequences (Nakano, T. and T. Graf (1991) Biochim. Biophys. Acta 1090:273-276). In chickens, both forms of lysozyme are found in neutrophil granulocytes (heterophils), but only chicken-type lysozyme is found in egg white.

Generally, chicken-type lysozyme mRNA is found in both adherent monocytes and macrophages and nonadherent promyelocytes and granulocytes as well as in cells of the bone marrow, spleen, bursa, and oviduct. Goose-type lysozyme mRNA is found in non-adherent cells of the bone marrow and lung. Several isozymes have been found in rabbits, including leukocytic, gastrointestinal, and possibly lymphoepithelial forms (OMIM #153450, supra; Nakano and Graf, supra; and GenBank GI 1310929). A human lysozyme gene encoding a protein similar to chicken-type lysozyme has been cloned (Yoshimura, K. et al. (1988) Biochem. Biophys. Res. Commun. 150:794-801). A consensus motif featuring regularly spaced cysteine residues has been derived from the lysozyme C enzymes of various species (Prosite PS00128, <http://expasy.hcuge.ch> Swiss Institute of Bioinformatics). Lysozyme C shares about 40% amino acid sequence identity with α -lactalbumin.

Lysozymes have several disease associations. Lysozymuria is observed in diabetic nephropathy (Shima, M. et al. (1986) Clin. Chem. 32:1818-1822), endemic nephropathy (Bruckner, I. et al. (1978) Med. Interne. 16:117-125), urinary tract infections (Heidegger, H. (1990) Minerva Ginecol. 42:243-250), and acute monocytic leukemia (Shaw, M.T. (1978) Am. J. Hematol. 4:97-103). Nakano (supra) suggested a role for lysozyme in host defense systems. Older rabbits with an

inherited lysozyme deficiency show increased susceptibility to infections, such as subcutaneous abscesses (OMIM #153450, supra). Human lysozyme gene mutations cause hereditary systemic amyloidosis, a rare autosomal dominant disease in which amyloid deposits form in the viscera, including the kidney, adrenal glands, spleen, and liver. This disease is usually fatal by the fifth decade. The amyloid deposits contain variant forms of lysozyme. Renal amyloidosis is the most common and potentially the most serious form of organ involvement (Pepys, M.B. et al. (1993) Nature 362:553-557; OMIM #105200 Familial Visceral Amyloidosis; Cotran, R.S. et al. (1994) Robbins Pathologic Basis of Disease, W.B. Saunders Company, Philadelphia PA, pp. 231-238). Increased levels of lysozyme and lactate have been observed in the cerebrospinal fluid of patients with bacterial meningitis (Ponka, A. et al. (1983) Infection 11:129-131). Acute monocytic leukemia is characterized by massive lysozymuria (Den Tandt, W.R. (1988) Int. J. Biochem. 20:713-719).

The discovery of new hydrolase proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative, autoimmune/inflammatory, neurological, renal, adrenal, and genetic disorders.

SUMMARY OF THE INVENTION

The invention features substantially purified polypeptides, hydrolase proteins, referred to collectively as "HYDRL" and individually as "HYDRL-1," "HYDRL-2," "HYDRL-3," "HYDRL-4," "HYDRL-5," "HYDRL-6," "HYDRL-7," "HYDRL-8," "HYDRL-9," "HYDRL-10," "HYDRL-11," "HYDRL-12," "HYDRL-13," "HYDRL-14," "HYDRL-15," and "HYDRL-16." In one aspect, the invention provides a substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-16 and fragments thereof. The invention also includes a polypeptide comprising an amino acid sequence that differs by one or more conservative amino acid substitutions from an amino acid sequence selected from the group consisting of SEQ ID NO:1-16.

The invention further provides a substantially purified variant having at least 90% amino acid identity to at least one of the amino acid sequences selected from the group consisting of SEQ ID NO:1-16 and fragments thereof. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-16 and fragments thereof. The invention also includes an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-16 and fragments thereof.

Additionally, the invention provides an isolated and purified polynucleotide which hybridizes

under stringent conditions to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-16 and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide encoding the polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1-16 and fragments thereof.

The invention also provides a method for detecting a polynucleotide in a sample containing nucleic acids, the method comprising the steps of: (a) hybridizing the complement of the polynucleotide sequence to at least one of the polynucleotides of the sample, thereby forming a hybridization complex; and (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a polynucleotide in the sample. In one aspect, the method further comprises amplifying the polynucleotide prior to hybridization.

The invention also provides an isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32 and fragments thereof. The invention further provides an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32 and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32 and fragments thereof.

The invention further provides an expression vector containing at least a fragment of the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-16. In another aspect, the expression vector is contained within a host cell.

The invention also provides a method for producing a polypeptide, the method comprising the steps of: (a) culturing the host cell containing an expression vector containing a polynucleotide of the invention under conditions suitable for the expression of the polypeptide; and (b) recovering the polypeptide from the host cell culture.

The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-16 and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention further includes a purified antibody which binds to a polypeptide selected from the group consisting of SEQ ID NO:1-16 and fragments thereof. The invention also provides a purified agonist and a purified antagonist to the polypeptide.

The invention also provides a method for treating or preventing a disorder associated with

decreased expression or activity of HYDRL, the method comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-16 and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention also provides a method for treating or preventing a disorder associated with increased expression or activity of HYDRL, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16 and fragments thereof.

BRIEF DESCRIPTION OF THE FIGURE AND TABLES

Figure 1 shows the amino acid sequence alignment between HYDRL-1 (Incyte Clone ID 2293764; SEQ ID NO:1), Colobus guereza lysozyme-c precursor (GI 1790927; SEQ ID NO:33), Colobus angolensis lysozyme-c precursor (GI 1790967; SEQ ID NO:34), and Nasalis larvatus lysozyme-c precursor (GI 1790984; SEQ ID NO:35), produced using the multisequence alignment program of LASERGENE software (DNASTAR, Madison WI).

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding HYDRL.

Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods, algorithms, and searchable databases used for analysis of HYDRL.

Table 3 shows selected fragments of each nucleic acid sequence; the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding HYDRL were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze HYDRL, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will

be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms “a,” “an,” and “the” include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to “a host cell” includes a plurality of such host cells, and a reference to “an antibody” is a
5 reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be
10 used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

15 DEFINITIONS

“HYDRL” refers to the amino acid sequences of substantially purified HYDRL obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term “agonist” refers to a molecule which intensifies or mimics the biological activity of
20 HYDRL. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of HYDRL either by directly interacting with HYDRL or by acting on components of the biological pathway in which HYDRL participates.

An “allelic variant” is an alternative form of the gene encoding HYDRL. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in
25 polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

“Altered” nucleic acid sequences encoding HYDRL include those sequences with deletions,
30 insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as HYDRL or a polypeptide with at least one functional characteristic of HYDRL. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding HYDRL, and improper or unexpected hybridization to allelic

variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding HYDRL. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent HYDRL. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of HYDRL is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine.

Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms “amino acid” and “amino acid sequence” refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where “amino acid sequence” is recited to refer to an amino acid sequence of a naturally occurring protein molecule, “amino acid sequence” and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

“Amplification” relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term “antagonist” refers to a molecule which inhibits or attenuates the biological activity of HYDRL. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of HYDRL either by directly interacting with HYDRL or by acting on components of the biological pathway in which HYDRL participates.

The term “antibody” refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, F(ab')₂, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind HYDRL polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition containing a nucleic acid sequence which is complementary to the "sense" strand of a specific nucleic acid sequence. Antisense molecules may be produced by any method including synthesis or transcription. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form duplexes and to block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic HYDRL, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" and "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acid strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding HYDRL or fragments of HYDRL may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been resequenced to

resolve uncalled bases, extended using the XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of one or more Incyte Clones and, in some cases, one or more public domain ESTs, using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

“Conservative amino acid substitutions” are those substitutions that, when made, least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

	Original Residue	Conservative Substitution
	Ala	Gly, Ser
	Arg	His, Lys
	Asn	Asp, Gln, His
15	Asp	Asn, Glu
	Cys	Ala, Ser
	Gln	Asn, Glu, His
	Glu	Asp, Gln, His
	Gly	Ala
20	His	Asn, Arg, Gln, Glu
	Ile	Leu, Val
	Leu	Ile, Val
	Lys	Arg, Gln, Glu
	Met	Leu, Ile
25	Phe	His, Met, Leu, Trp, Tyr
	Ser	Cys, Thr
	Thr	Ser, Val
	Trp	Phe, Tyr
	Tyr	His, Phe, Trp
30	Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A “deletion” refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term “derivative” refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative

polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "fragment" is a unique portion of HYDRL or the polynucleotide encoding HYDRL which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50% of a polypeptide) as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:17-32 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:17-32, for example, as distinct from any other sequence in the same genome. A fragment of SEQ ID NO:17-32 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:17-32 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:17-32 and the region of SEQ ID NO:17-32 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-16 is encoded by a fragment of SEQ ID NO:17-32. A fragment of SEQ ID NO:1-16 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-16. For example, a fragment of SEQ ID NO:1-16 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-16. The precise length of a fragment of SEQ ID NO:1-16 and the region of SEQ ID NO:1-16 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

The term “similarity” refers to a degree of complementarity. There may be partial similarity or complete similarity. The word “identity” may substitute for the word “similarity.” A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as “substantially similar.” The inhibition of hybridization of the

“BLAST 2 Sequences” tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the “BLAST 2 Sequences” tool Version 2.0.9 (May-07-1999) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases “percent identity” and “% identity,” as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the hydrophobicity and acidity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and “diagonals saved”=5. The PAM250 matrix is selected as the default

residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

The term "humanized antibody" refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of identity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive

annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Generally, such wash temperatures are selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

The term "microarray" refers to an arrangement of distinct polynucleotides on a substrate.

The terms “element” and “array element” in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

The term “modulate” refers to a change in the activity of HYDRL. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of HYDRL.

The phrases “nucleic acid” and “nucleic acid sequence” refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

“Peptide nucleic acid” (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Probe" refers to nucleic acid sequences encoding HYDRL, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for

recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be used to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

The term “sample” is used in its broadest sense. A sample suspected of containing nucleic acids encoding HYDRL, or fragments thereof, or HYDRL itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms “specific binding” and “specifically binding” refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope “A,” the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term “substantially purified” refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably about 75% free, and most preferably about 90% free from other components with which they are naturally associated.

A “substitution” refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

“Substrate” refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

“Transformation” describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term “transformed” cells includes stably transformed cells in which the inserted DNA is capable of replication either as an

autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human hydrolase proteins (HYDRL), the polynucleotides encoding HYDRL, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative, autoimmune/inflammatory, neurological, renal, adrenal, and genetic disorders.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding HYDRL. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each HYDRL were identified, and column 4 shows the cDNA

libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. The Incyte clones in column 5 were used to assemble the consensus nucleotide sequence of each HYDRL and are useful as fragments in hybridization technologies.

5 The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; column 6 shows homologous sequences as identified by BLAST analysis; and column 7 shows analytical
10 methods and in some cases, searchable databases to which the analytical methods were applied. The methods of column 7 were used to characterize each polypeptide through sequence homology and protein motifs.

As shown in Figure 1, HYDRL-1 has chemical and structural similarity with Colobus guereza lysozyme-c precursor (GI 1790927; SEQ ID NO:33), Colobus angolensis lysozyme-c precursor (GI
15 1790967; SEQ ID NO:34) and Nasalis larvatis lysozyme-c precursor (GI 1790984; SEQ ID NO:35). In particular, HYDRL-1 and Colobus guereza lysozyme-c precursor share 40% identity, HYDRL-1 and Colobus angolensis lysozyme-c precursor share 40% identity, and HYDRL-1 and Nasalis larvatis lysozyme-c precursor share 41% identity.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions
20 associated with nucleotide sequences encoding HYDRL. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists fragments of the nucleotide sequences of column 1. These fragments are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:17-32 and to distinguish between SEQ ID NO:17-32 and related polynucleotide sequences. The polypeptides encoded by these fragments are useful, for example, as immunogenic peptides. Column
25 3 lists tissue categories which express HYDRL as a fraction of total tissues expressing HYDRL. Column 4 lists diseases, disorders, or conditions associated with those tissues expressing HYDRL as a fraction of total tissues expressing HYDRL. Column 5 lists the vectors used to subclone each cDNA library.

Northern analysis of SEQ ID NO:17 shows the expression of this sequence in tissue
30 associated with cancer.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding HYDRL were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

SEQ ID NO:18 maps to chromosome 6 within the interval from 42.30 to 45.40 centiMorgans, to chromosome 9 within the interval from 130.40 to 166.50 centiMorgans, and to chromosome 16 within the interval from 88.10 to 92.60 centiMorgans.

SEQ ID NO:25 maps to chromosome 1 within the interval from 22.90 to 39.90 centiMorgans and to chromosome 3 within the interval from 30.90 to 43.00 centiMorgans. The interval on chromosome 3 from 30.90 to 43.00 centiMorgans also contains an EST associated with von Hippel-Lindau syndrome.

SEQ ID NO:28 maps to chromosome 10 within the interval from 137.60 to 139.20 centiMorgans.

The invention also encompasses HYDRL variants. A preferred HYDRL variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the HYDRL amino acid sequence, and which contains at least one functional or structural characteristic of HYDRL.

The invention also encompasses polynucleotides which encode HYDRL. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:17-32, which encodes HYDRL.

The invention also encompasses a variant of a polynucleotide sequence encoding HYDRL. In particular, such a variant polynucleotide sequence will have at least about 80%, or alternatively at least about 90%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding HYDRL. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:17-32 which has at least about 80%, or alternatively at least about 90%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:17-32. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of HYDRL.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding HYDRL, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring HYDRL, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode HYDRL and its variants are generally capable

of hybridizing to the nucleotide sequence of the naturally occurring HYDRL under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding HYDRL or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding HYDRL and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode HYDRL and HYDRL derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding HYDRL or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:17-32 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Perkin-Elmer). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Perkin-Elmer), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding HYDRL may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.)

Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060.)

Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be

present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode HYDRL may be cloned in recombinant DNA molecules that direct expression of HYDRL, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express HYDRL.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter HYDRL-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

In another embodiment, sequences encoding HYDRL may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) *Nucleic Acids Symp. Ser.* 7:215-223; and Horn, T. et al. (1980) *Nucleic Acids Symp. Ser.* 7:225-232.) Alternatively, HYDRL itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) *Science* 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of HYDRL, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) *Methods Enzymol.* 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY.)

In order to express a biologically active HYDRL, the nucleotide sequences encoding HYDRL or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding HYDRL. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of

sequences encoding HYDRL. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding HYDRL and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding HYDRL and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding HYDRL. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding HYDRL. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding HYDRL can be achieved using a multifunctional *E. coli* vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding HYDRL into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for *in vitro* transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of HYDRL are needed, e.g. for the production of antibodies, vectors which direct high level expression of HYDRL may be used. For example, vectors

containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of HYDRL. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such
 5 vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; and Scorer, C.A. et al. (1994) *Bio/Technology* 12:181-184.)

Plant systems may also be used for expression of HYDRL. Transcription of sequences
 10 encoding HYDRL may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105.)
 15 These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding HYDRL may be ligated into
 20 an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses HYDRL in host cells. (See, e.g., Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-
 25 based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) *Nat. Genet.*
 30 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of HYDRL in cell lines is preferred. For example, sequences encoding HYDRL can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the

introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *apr* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β glucuronidase and its substrate β -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding HYDRL is inserted within a marker gene sequence, transformed cells containing sequences encoding HYDRL can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding HYDRL under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding HYDRL and that express HYDRL may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of HYDRL using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques

include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on HYDRL is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding HYDRL include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding HYDRL, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding HYDRL may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode HYDRL may be designed to contain signal sequences which direct secretion of HYDRL through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct

modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding HYDRL may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric HYDRL protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of HYDRL activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the HYDRL encoding sequence and the heterologous protein sequence, so that HYDRL may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled HYDRL may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ³⁵S-methionine.

Fragments of HYDRL may be produced not only by recombinant means, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra, pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A peptide synthesizer (Perkin-Elmer). Various fragments of HYDRL may be synthesized separately and then combined to produce the full length molecule.

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of HYDRL and hydrolase proteins. In addition, the expression of HYDRL is closely associated with proliferating tissues, inflamed tissues, neurological tissues, and cancer. In some cases, sequences encoding HYDRL map to chromosomal regions associated with inherited diseases.

Therefore, HYDRL appears to play a role in cell proliferative, autoimmune/inflammatory, neurological, renal, adrenal, and genetic disorders. In the treatment of disorders associated with increased HYDRL expression or activity, it is desirable to decrease the expression or activity of HYDRL. In the treatment of disorders associated with decreased HYDRL expression or activity, it is desirable to increase the expression or activity of HYDRL.

Therefore, in one embodiment, HYDRL or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HYDRL. Examples of such disorders include, but are not limited to, a cell proliferative disorder, such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and a cancer including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder, such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation; viral, bacterial, fungal, parasitic, protozoal, and helminthic infections; trauma, and hematopoietic cancer including lymphoma, leukemia, and myeloma; a neurological disorder, such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders,

trifunctional protein deficiency, and mitochondrial short-chain 3-hydroxyacyl-CoA dehydrogenase deficiency.

In another embodiment, a vector capable of expressing HYDRL or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HYDRL including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified HYDRL in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HYDRL including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of HYDRL may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HYDRL including, but not limited to, those listed above.

In a further embodiment, an antagonist of HYDRL may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HYDRL. Examples of such disorders include, but are not limited to, those cell proliferative, autoimmune/inflammatory, neurological, renal, adrenal, and genetic disorders described above. In one aspect, an antibody which specifically binds HYDRL may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express HYDRL.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding HYDRL may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HYDRL including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of HYDRL may be produced using methods which are generally known in the art. In particular, purified HYDRL may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind HYDRL. Antibodies to HYDRL may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit

dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with HYDRL or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to HYDRL have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of HYDRL amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to HYDRL may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce HYDRL-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding HYDRL, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding HYDRL may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding HYDRL. Thus, complementary molecules or fragments may be used to modulate HYDRL activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding HYDRL.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding HYDRL. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

Genes encoding HYDRL can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding HYDRL. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding HYDRL. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may be employed. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of

RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding HYDRL.

5 Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of
10 candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis.

15 Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding HYDRL. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

20 RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine,
25 and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient.

30 Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and

monkeys.

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of HYDRL, antibodies to HYDRL, and mimetics, agonists, antagonists, or inhibitors of HYDRL. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures.

determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example HYDRL or fragments thereof, antibodies of HYDRL, and agonists, antagonists or inhibitors of HYDRL, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg , up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art.

Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind HYDRL may be used for the diagnosis of disorders characterized by expression of HYDRL, or in assays to monitor patients being treated with HYDRL or agonists, antagonists, or inhibitors of HYDRL. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for HYDRL include methods which utilize the antibody and a label to detect HYDRL in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring HYDRL, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of HYDRL expression. Normal or standard values for HYDRL expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibody to HYDRL under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of HYDRL expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding HYDRL may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of HYDRL may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of HYDRL, and to monitor regulation of HYDRL levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding HYDRL or closely related molecules may be used to identify nucleic acid sequences which encode HYDRL. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding HYDRL, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the HYDRL encoding sequences. The hybridization probes of the subject

invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:17-32 or from genomic sequences including promoters, enhancers, and introns of the HYDRL gene.

Means for producing specific hybridization probes for DNAs encoding HYDRL include the cloning of polynucleotide sequences encoding HYDRL or HYDRL derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S , or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding HYDRL may be used for the diagnosis of disorders associated with expression of HYDRL. Examples of such disorders include, but are not limited to, a cell proliferative disorder, such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and a cancer including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder, such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation; viral, bacterial, fungal, parasitic, protozoal, and helminthic infections;

palmitoyl transferase and carnitine deficiency, mitochondrial very-long-chain acyl-CoA dehydrogenase deficiency, mitochondrial medium-chain acyl-CoA dehydrogenase deficiency, mitochondrial short-chain acyl-CoA dehydrogenase deficiency, mitochondrial electron transport flavoprotein and electron transport flavoprotein:ubiquinone oxidoreductase deficiency, mitochondrial trifunctional protein deficiency, and mitochondrial short-chain 3-hydroxyacyl-CoA dehydrogenase deficiency. The polynucleotide sequences encoding HYDRL may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered HYDRL expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding HYDRL may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding HYDRL may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding HYDRL in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of HYDRL, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding HYDRL, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or

overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding HYDRL may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding HYDRL, or a fragment of a polynucleotide complementary to the polynucleotide encoding HYDRL, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

Methods which may also be used to quantify the expression of HYDRL include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) *J. Immunol. Methods* 159:235-244; Duplaa, C. et al. (1993) *Anal. Biochem.* 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding HYDRL may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial

chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding HYDRL on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder.

The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, HYDRL, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between HYDRL and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with HYDRL, or fragments thereof, and washed. Bound HYDRL is then detected by methods well known in the art. Purified HYDRL can also be coated directly onto plates for use in the aforementioned drug screening techniques.

Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding HYDRL specifically compete with a test compound for binding
5 HYDRL. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with HYDRL.

In additional embodiments, the nucleotide sequences which encode HYDRL may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such
10 properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

15 The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. [Attorney Docket No. PF-0634 P, filed November 12, 1998] and U.S. Ser. No. 60/135,519, are hereby expressly incorporated by reference.

EXAMPLES

20 I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged
25 over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A⁺) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN,
30 Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP

vector system (Stratagene) or SUPERScript plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSORT1 plasmid (Life Technologies), or pINCY (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

cDNA sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Perkin-Elmer) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:17-32. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

5 IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

10 Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$15 \quad \frac{\% \text{ sequence identity} \times \% \text{ maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are
20 usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding HYDRL occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic,
25 developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in
30 Table 3.

V. Chromosomal Mapping of HYDRL Encoding Polynucleotides

The cDNA sequences which were used to assemble SEQ ID NO:18-32 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched

SEQ ID NO:18-32 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 5). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO., to that map location.

The genetic map locations of SEQ ID NO:18, SEQ ID NO:25, and SEQ ID NO:28 are described in The Invention as ranges, or intervals, of human chromosomes. More than one map location is reported for SEQ ID NO:18 and SEQ ID NO:25, indicating that previously mapped sequences having similarity, but not complete identity, to SEQ ID NO:18 and SEQ ID NO:25 were assembled into their respective clusters. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Diseases associated with the public and Incyte sequences located within the indicated intervals are also reported in the Invention where applicable.

VI. Extension of HYDRL Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:17-32 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+} , $(NH_4)_2SO_4$, and β -mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer

pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 µl to 10 µl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

In like manner, the nucleotide sequences of SEQ ID NO:17-32 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an

appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

IX. Complementary Polynucleotides

Sequences complementary to the HYDRL-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring HYDRL. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of HYDRL. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the HYDRL-encoding transcript.

X. Expression of HYDRL

Expression and purification of HYDRL is achieved using bacterial or virus-based expression systems. For expression of HYDRL in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express HYDRL upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of HYDRL in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding HYDRL by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, HYDRL is synthesized as a fusion protein with, e.g., glutathione

S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from HYDRL at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified HYDRL obtained by these methods can be used directly in the following activity assay.

XI. Demonstration of HYDRL Activity

For purposes of example, assays measuring the β -glucosidase activity and the lysozyme activity of an HYDRL molecule are described. In a β -glucosidase activity assay, varying amounts of HYDRL are incubated with 1 mM 4-nitrophenyl β -D-glycopyranoside (a substrate) in 50 mM sodium acetate buffer, pH 5.0, for various times (typically 1-5 minutes) at 37°C. The reaction is halted by heating to 100°C for 2 minutes. The absorbance is measured spectrophotometrically at 410 nm, and is proportional to the β -glucosidase activity of HYDRL in the sample. (See, e.g., Hrmova, M. et al. (1998) J. Biol. Chem. 273:11134-11143.)

Lysozyme activity of HYDRL is demonstrated by its ability to lyse Micrococcus lysodeikticus bacterial cells. (See, e.g., Enzymatic Assay of Lysozyme 1, Sigma Aldrich, St. Louis MO). A 0.015% suspension of lyophilized Micrococcus lysodeikticus cells (ATCC 4698) is prepared in 66 mM potassium phosphate buffer, pH 6.24 (Buffer A) at 25°C. A 2.5 ml aliquot of the cell suspension is pipetted into a optical cuvette and equilibrated to 25°C. The absorbance at 450 nm is monitored until constant, between 0.6 and 0.7, using a thermostatted spectrophotometer. A blank reaction is prepared in a second cuvette containing 2.5 ml Buffer A. HYDRL is dissolved in cold Buffer A. A 0.1 ml aliquot of the HYDRL solution is added to the test cuvette, and 0.1 ml Buffer A is added to the blank cuvette. The cuvettes are immediately mixed by inversion, and the decrease in absorbance at 450 nm is recorded for approximately 5 minutes. As the bacteria lyse, the turbidity of the solution, and hence the absorbance at 450 nm, decrease. The rate of the decrease in absorbance at 450 nm in the test cuvette is proportional to the lysozyme activity of HYDRL in the original sample.

XII. Functional Assays

HYDRL function is assessed by expressing the sequences encoding HYDRL at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a

mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT (Life Technologies) and pCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μ g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of HYDRL on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding HYDRL and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding HYDRL and other genes of interest can be analyzed by northern analysis or microarray techniques.

XIII. Production of HYDRL Specific Antibodies

HYDRL substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the HYDRL amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for

selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-HYDRL activity by, for example, binding the peptide or HYDRL to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIV. Purification of Naturally Occurring HYDRL Using Specific Antibodies

Naturally occurring or recombinant HYDRL is substantially purified by immunoaffinity chromatography using antibodies specific for HYDRL. An immunoaffinity column is constructed by covalently coupling anti-HYDRL antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing HYDRL are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of HYDRL (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/HYDRL binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and HYDRL is collected.

XV. Identification of Molecules Which Interact with HYDRL

HYDRL, or biologically active fragments thereof, are labeled with ¹²⁵I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled HYDRL, washed, and any wells with labeled HYDRL complex are assayed. Data obtained using different concentrations of HYDRL are used to calculate values for the number, affinity, and association of HYDRL with the candidate molecules.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are

obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	17	2293764	BRAINON01	2293764H1 (BRAINON01), 2293764R6 (BRAINON01), 2293764T6 (BRAINON01)
2	18	949738	PANCNOT05	938735H1 (CERVNOT01), 938735X13 (CERVNOT01), 949738H1 (PANCNOT05), 1500746F1 (SINTBST01), 2206160F6 (SPLNFET02)
3	19	1297034	BRSTNOT07	1297034H1 (BRSTNOT07), 1713735F6 (UCMCNOT02), 1713735T6 (UCMCNOT02), 1889362F6 (BLADTUT07), 822634R1 (KERANOT02), 903616R2 (COLNNOT07), SAQA00517F1, SAQA00539F1
4	20	1553276	BLADTUT04	1553276F6 (BLADTUT04), 1553276H1 (BLADTUT04)
5	21	1702211	BLADTUT05	1702211H1 (BLADTUT05), 2768772F6 (COLANOT02), SCCA04943V1, SCCA02992V1, SCCA02905V1, SCCA03592V1, SCCA05380V1
6	22	1859618	PROSNOT18	168238H1 (LIVRNOT01), 294798R6 (LIVRNOT04), 1859618F6 (PROSNOT18), 1859618H1 (PROSNOT18), SARB00217F1, SARB00638F1, SARB00588F1
7	23	2011071	TESTNOT03	2011071H1 (TESTNOT03), 2011071R6 (TESTNOT03), 2011071T6 (TESTNOT03)
8	24	2186517	PROSNOT26	1286776F1 (BRAINOT11), 2186517H1 (PROSNOT26), 2186517X13C1 (PROSNOT26), 2465845F6 (THYRNOT08), 2615896F6 (GBLANOT01), 3250090H1 (SEMVNOT03), 4820270F6 (PROSTUT17)
9	25	2253585	OVARTUT01	1361776F1 (LUNGNOT12), 1686637T6 (PROSNOT15), 2253585H1 (OVARTUT01), 2822491F6 (ADRETUT06), 2822491T6 (ADRETUT06)
10	26	2447520	THPINOT03	079381F1 (SYNORAB01), 1616379H1 (BRAITUT12), 2395202H1 (THPIAZT01), 2447520H1 (THPINOT03), 2682142H1 (SINIUCT01), 2848332H1 (BRSTTUT13), 2856151H1 (CONNNOT02), 3595358H1 (FIBPNOT01), 3702357H1 (PENCNOT07), 3706925H1 (PENCNOT07), 3919689H1 (BRAINOT14)

Table 1 (cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
11	27	2481345	SMCANOT01	2481345F6 (SMCANOT01), 2481345H1 (SMCANOT01), 4099083H2 (BRAITUT26), 4826226H1 (BLADDIT01)
12	28	2484020	SMCANOT01	079652F1 (SYNORAB01), 2308533H1 (NGANNOT01), 2484020H1 (SMCANOT01), 2799586H1 (PENCNOT01), 3571038H1 (HEAPNOT01)
13	29	2862528	SININOT03	403408R1 (TMLR3DT01), 875946R1 (LUNGAST01), 928587R1 (BRAINOT04), 2135155H1 (ENDCNOT01), 2862528H1 (SININOT03)
14	30	3200650	PENCNOT02	1441391F1 (THYRNOT03), 1580635F6 (DUODNOT01), 1726450F6 (PROSNOT14), 2225236T6 (SEMVNOT01), 2995844F6 (OVRTUT07), 3200650F6 (PENCNOT02), 3200650H1 (PENCNOT02), 3319055F6 (PROSBPT03), 3325495H2 (PTHYNOT03), 3358591H1 (PROSTUT16), 3589008H1 (293TF5T01), 4823564H1 (PROSTUT17)
15	31	4107621	BRSTTUT17	1693538F6 (COLNNOT23), 2125414F6 (BRSTNOT07), 2853982H1 (CONNNOT02), 2967492F6 (SCORNOT04), 2967492T6 (SCORNOT04), 4107621H1 (BRSTTUT17), 4648345H1 (PROSTUT20)
16	32	4661133	BRAVXTUT03	487982R1 (HNT2AGT01), 1553195X11 (BLADTUT04), 2748561H1 (LUNGUTUT11), 4661133H1 (BRAVXTUT03), SCAA03254V1, SCAA05351V1, SCAA00669V1

Table 2

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification/ Homologies	Analytical Methods
1	159	S57 S136 S130		Alpha-lactalbumin/lysozyme C motifs: Y24-E34, G37-T61, Y44-A53, D69-D110, S71-N87, N94-V119, C114-C135, G124-C135 Lactalbumin/lysozyme consensus sequence: C96-C114 Lysyzome HMM motif: K22-C147 Lysozyme signature motifs: I23-Y41, H54-H66, T90-K116, N123-G146 Signal peptide: M1-A21	Lysozyme [Paralichthys olivaceus] (GeneSeq W69514) Lysozyme c precursor [Colobus guereza] (GI 1790927) Lysozyme c precursor [Colobus angolensis] (GI 1790967) Lysozyme c precursor [Nasalis larvatis] (GI 1790984)	BLAST BLOCKS MOTIFS PFAM PRINTS SPScan
2	285	T3 T132 T155 T211 S261 S263 S280	N121		NG,NG-dimethylarginine dimethylaminohydrolase [Rattus norvegicus] (GI 1906800)	BLAST MOTIFS
3	331	S290 T94 T156 T216 T303 S308 S322 S52 T123 T225	N168 N198	Protein phosphodiesterases domain: R64-S74, P78-D113 Signal peptide: M1-R26	Similarity to E. coli glycerophosphoryldiester phosphodiesterase [C. elegans] (GI 3877620)	BLAST HMM MOTIFS PRODOM SPScan

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification/ Homologies	Analytical Methods
4	153	T31		Signal peptide: M1-S21	Colipase [Myocastor coyus] (GI 599867)	BLAST HMM MOTIFS SPScan
5	571	T27 S119 S131 T158 S187 T188 S285 S324 S366 T379 T390 S398 T486 T53 S106 S139 S304 S477	N105	Carboxylesterase domain: R6-W547 Signal peptide: M1-T27	Carboxylesterase [Mus musculus] (GI 404389)	BLAST BLOCKS HMM MOTIFS PFAM ProfileScan SPScan
6	347	S39 S274 S323	N79 N186 N269 N306 N325	Leucine rich repeat domains: K93-A140, T141-T188, L189-P236, D237-G284 Signal peptide: M1-A35	Phospholipase A2 inhibitor [Agkistrodon blomhoffii siniticus] (GI 3358089)	BLAST BLOCKS MOTIFS PFAM PRINTS SPScan
7	194	S18 T61 T109 T23	N104	C-type lysozyme/alpha-lactalbumin family domain: K66-C191	Lysozyme c precursor [Erythrocebus patas] (GI 1791001)	BLAST BLOCKS MOTIFS Pfam PRINTS ProfileScan

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification/ Homologies	Analytical Methods
8	361	T304 S172 S249 T308 T337 S29 S47 T147 S161 Y170	N106	Metallo-beta-lactamase family domain: V97-H267 Signal peptide: M1-A35	Glyoxalase II [Homo sapiens] (GI 1237213)	BLAST BLOCKS MOTIFS PFAM SPScan
9	306	S158 S47 T84		Alpha/beta hydrolase fold: R86-L305 ATP/GTP-binding site motif A (P-loop): G130-T137	Similar to alpha/beta hydrolase fold [C. elegans] (GI 3878848)	BLAST MOTIFS PFAM PRINTS
10	483	S99 T118 T185 S279 S360 S2 S37 T95 T326	N317		Mitochondrial very-long-chain acyl-CoA thioesterase [Rattus norvegicus] (GI 2832739)	BLAST MOTIFS
11	144	T111 T123 S139 Y109	N108	Sulfatases protein domains: P32-F48, P80-L91, G120-D130 Signal Peptide: M1-G22	Bone-related sulphatase-like precursor [Mus musculus] (GeneSeq R51355)	BLAST BLOCKS HMM MOTIFS SPScan
12	180	S28 T99 S158 S39 S178	N7 N153	mutT domain: V34-V73 MutT domain signature: G51-E70	Diphosphoinositol polyphosphate phosphohydrolase [Homo sapiens] (GI 3978224)	BLAST BLOCKS MOTIFS PFAM PRINTS

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification/ Homologies	Analytical Methods
13	375	T145 S34 S172 T312 S78 T121 S156 T211 S254 S261 S280 S340 S362	N262	Signal peptide: M1-A29	Acyl-CoA thioester hydrolase [Homo sapiens] (GI 1906670)	BLAST HMM MOTIFS SPScan
14	637	S142 S254 T341 S396 S541 T561 S5 T117 T164 T356 S410 T468		Glycosyl hydrolase family domain: A11-H629 Signal peptide: M1-R28	Beta-galactosidase [Bacillus circulans] (GI 2289790)	BLAST BLOCKS HMM MOTIFS PFAM PRINTS SPScan
15	314	S46 T120 S257 T272 T148 S171 T199 S232 T268	N44 N255	Alpha/beta hydrolase fold: F59-L304	Kraken [Drosophila melanogaster] (GI 2274926)	BLAST MOTIFS PFAM PRINTS
16	448	S95 S105 T186 S232 S263 T269 S276 S388 S406 T417 T119 S145 S284 T439 Y424			Acyl-CoA thioesterase [Mus musculus] (GI 5102774)	BLAST MOTIFS

05021455, 0502001

Table 3

Nucleotide SEQ ID NO:	Selected Fragments of Nucleotide Sequence	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
17	168-212		Cancer	
18	323-373 1067-1117	Reproductive (0.280) Nervous (0.189) Gastrointestinal (0.122) Cardiovascular (0.104)	Cancer (0.476) Inflammation (0.220) Cell Proliferation (0.177) Trauma (0.110)	PSPORT1
19	161-220 668-727	Reproductive (0.258) Nervous (0.175) Gastrointestinal (0.155) Cardiovascular (0.113)	Cancer (0.495) Inflammation (0.186) Cell Proliferation (0.175)	pINCY
20	68-127 260-304 401-448	Urologic (1.000)	Cancer (1.000)	pINCY
21	1680-1736	Gastrointestinal (0.600) Hematopoietic/Immune (0.300)	Inflammation (0.500) Cancer (0.400) Cell Proliferation (0.200)	pINCY
22	179-238	Gastrointestinal (0.478) Reproductive (0.217) Hematopoietic/Immune (0.130) Cardiovascular (0.130)	Cancer (0.435) Inflammation (0.217)	pINCY
23	61-120	Reproductive (1.000)	Inflammation (1.000)	PBLUESCRIPT
24	247-297 811-870	Gastrointestinal (0.312) Reproductive (0.312)	Cancer (0.438) Inflammation (0.250)	pINCY
25	86-145	Gastrointestinal (0.308) Reproductive (0.308) Hematopoietic/Immune (0.115) Nervous (0.115)	Cancer (0.692)	PSPORT1

Table 3 (cont.)

Nucleotide SEQ ID NO:	Selected Fragments of Nucleotide Sequence	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
26	326-385 817-874	Reproductive (0.225) Nervous (0.169) Gastrointestinal (0.169) Cardiovascular (0.113)	Cancer (0.423) Inflammation (0.254) Cell Proliferation (0.155) Trauma (0.127)	pINCY
27	406-465 520-564	Cardiovascular (0.500) Nervous (0.500)	Cancer (0.500) Cell proliferation (0.500)	pINCY
28	516-575	Gastrointestinal (0.200) Cardiovascular (0.171) Reproductive (0.171) Nervous (0.143)	Cancer (0.429) Inflammation (0.257) Cell Proliferation (0.143) Trauma (0.114)	pINCY
29	986-1045	Nervous (0.229) Reproductive (0.186) Gastrointestinal (0.143) Cardiovascular (0.100) Musculoskeletal (0.100)	Cancer (0.400) Cell Proliferation (0.286) Inflammation (0.214)	pINCY
30	1-58	Reproductive (0.500) Gastrointestinal (0.206) Nervous (0.176)	Cancer (0.765)	pINCY
31	115-174	Reproductive (0.375) Hematopoietic/Immune (0.156) Nervous (0.156) Gastrointestinal (0.125)	Cancer (0.562) Inflammation (0.312)	pINCY
32	1-146	Cardiovascular (0.312) Reproductive (0.292) Nervous (0.125)	Cancer (0.458) Cell Proliferation (0.229) Inflammation (0.188)	pINCY

Table 4

Nucleotide SEQ ID NO:	Library	Library Comment
17	BRAINOT01	Library was constructed and normalized from 4.88 million independent clones from the BRAINOT03 brain tissue library. RNA was made from brain tissue removed from a 26-year-old Caucasian male during cranioplasty and excision of a cerebral meningeal lesion. Pathology for the associated tumor tissue indicated a grade 4 oligoastrocytoma in the right frontoparietal part of the brain. 4.9×10^6 independent clones of the plasmid library in <i>E. coli</i> strain DH12S (Life Technologies) were grown in liquid culture and then superinfected with a 5-fold excess of the helper phage M13K07. To reduce the number of excess cDNA copies according to their abundance levels in the library, the cDNA library was normalized (Soares et al. (1994) Proc. Natl. Acad. Sci. USA 91:9928-9932).
18	PANCNOT05	Library was constructed using RNA isolated from the pancreatic tissue of a 2-year-old Hispanic male who died from cerebral anoxia.
19	BRSTNOT07	Library was constructed using RNA isolated from diseased breast tissue removed from a 43-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated mildly proliferative fibrocystic changes with epithelial hyperplasia, papillomatosis, and duct ectasia. Pathology for the associated tumor tissue indicated invasive grade 4, nuclear grade 3 mammary adenocarcinoma with extensive comedo necrosis. Family history included epilepsy, cardiovascular disease, and type II diabetes.
20	BLADTUT04	Library was constructed using RNA isolated from bladder tumor tissue removed from a 60-year-old Caucasian male during a radical cystectomy, prostatectomy, and vasectomy. Pathology indicated grade 3 transitional cell carcinoma in the left bladder wall. Carcinoma in-situ was identified in the dome and trigone. Patient history included tobacco use. Family history included type I diabetes, a malignant neoplasm of the stomach, atherosclerotic coronary artery disease, and an acute myocardial infarction.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Comment
21	BLADTUT05	Library was constructed using RNA isolated from bladder tumor tissue removed from a 66-year-old Caucasian male during a radical prostatectomy, radical cystectomy, and urinary diversion. Pathology indicated grade 3 transitional cell carcinoma on the anterior wall of the bladder. Patient history included lung neoplasm and tobacco abuse in remission. Family history included a malignant breast neoplasm, tuberculosis, cerebrovascular disease, atherosclerotic coronary artery disease, and lung cancer.
22	PROSNOT18	Library was constructed using RNA isolated from diseased prostate tissue removed from a 58-year-old Caucasian male during a radical cystectomy, radical prostatectomy, and gastrectomy. Pathology indicated adenofibromatous hyperplasia associated with a grade 3 transitional cell carcinoma. Patient history included angina and emphysema. Family history included acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.
23	TESTNOT03	Library was constructed using RNA isolated from testicular tissue removed from a 37-year-old Caucasian male, who died from liver disease. Patient history included cirrhosis, jaundice, and liver failure.
24	PROSNOT26	Library was constructed using RNA isolated from prostate tissue removed from a 65-year-old Caucasian male during a radical prostatectomy. Pathology for the associated tumor tissue indicated an adenocarcinoma. The patient presented with elevated prostate specific antigen. Family history included a malignant stomach neoplasm.
25	OVRTUT01	Library was constructed using RNA isolated from ovarian tumor tissue removed from a 43-year-old Caucasian female during removal of the fallopian tubes and ovaries. Pathology indicated grade 2 mucinous cystadenocarcinoma involving the entire left ovary. Patient history included mitral valve disorder, pneumonia, and viral hepatitis. Family history included atherosclerotic coronary artery disease, pancreatic cancer, stress reaction, cerebrovascular disease, breast cancer, and uterine cancer.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Comment
26	THPINOT03	Library was constructed using RNA isolated from untreated THP-1 cells (ATCC TIB 202). THP-1 is a human promonocyte line derived from the peripheral blood of a 1-year-old Caucasian male with acute monocytic leukemia.
27	SMCANOT01	Library was constructed using RNA isolated from an aortic smooth muscle cell line derived from the explanted heart of a male during a heart transplant.
28	SMCANOT01	Library was constructed using RNA isolated from an aortic smooth muscle cell line derived from the explanted heart of a male during a heart transplant.
29	SININOT03	Library was constructed using RNA isolated from ileum tissue obtained from an 8-year-old Caucasian female, who died from head trauma. Serology was positive for cytomegalovirus.
30	PENCNOT02	Library was constructed using RNA isolated from penis right corpus cavernosum tissue.
31	BRSTTUT17	Library was constructed using RNA isolated from left breast tumor tissue removed from a 65-year-old Caucasian female during a unilateral radical mastectomy. Pathology indicated invasive and in situ grade 3, nuclear grade 2 ductal carcinoma. Patient history included hyperlipidemia and uterine leiomyoma. Family history included stomach cancer, myocardial infarction, atherosclerotic coronary artery disease, prostate cancer, benign hypertension, breast cancer, and hyperlipidemia.
32	BRVXTXT03	Library was constructed using RNA isolated from treated astrocytes removed from the brain of a female fetus who died after 22 weeks' gestation. The cells were treated with tumor necrosis factor alpha and interleukin 1, 10ng/ml each, for 24 hours.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value=1.06E-6 Assembled ESTs: fasta Identity= 95% or greater and Match length=200 bases or greater, fastx E value=1.0E-8 or less Full Length sequences: fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families

Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Normalized quality score \geq GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. <i>supra</i> ; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. A substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, and fragments thereof.

2. A substantially purified variant having at least 90% amino acid sequence identity to the amino acid sequence of claim 1.

3. An isolated and purified polynucleotide encoding the polypeptide of claim 1.

4. An isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide of claim 3.

5. An isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide of claim 3.

6. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 3.

7. A method for detecting a polynucleotide, the method comprising the steps of:
 (a) hybridizing the polynucleotide of claim 6 to at least one nucleic acid in a sample, thereby forming a hybridization complex; and
 (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of the polynucleotide in the sample.

8. The method of claim 7 further comprising amplifying the polynucleotide prior to hybridization.

9. An isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32 and fragments thereof.

10. An isolated and purified polynucleotide variant having at least 90% polynucleotide

sequence identity to the polynucleotide of claim 9.

11. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 9.

5

12. An expression vector comprising at least a fragment of the polynucleotide of claim 3.

13. A host cell comprising the expression vector of claim 12.

10

14. A method for producing a polypeptide, the method comprising the steps of:

a) culturing the host cell of claim 13 under conditions suitable for the expression of the polypeptide; and

b) recovering the polypeptide from the host cell culture.

15

15. A pharmaceutical composition comprising the polypeptide of claim 1 in conjunction with a suitable pharmaceutical carrier.

16. A purified antibody which specifically binds to the polypeptide of claim 1.

20

17. A purified agonist of the polypeptide of claim 1.

18. A purified antagonist of the polypeptide of claim 1.

25

19. A method for treating or preventing a disorder associated with decreased expression or activity of HYDRL, the method comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 15.

30

20. A method for treating or preventing a disorder associated with increased expression or activity of HYDRL, the method comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 18.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C12N 15/55, 9/14, C12Q 1/68, C12N 1/21, A61K 38/46, C07K 16/40		A3	(11) International Publication Number: WO 00/28045
			(43) International Publication Date: 18 May 2000 (18.05.00)
(21) International Application Number: PCT/US99/27009		2382 Lass Drive, Santa Clara, CA 95054 (US). <u>BANDMAN</u> , Olga [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). <u>CORLEY</u> , Neil, C. [US/US]; 1240 Dale Avenue, #30, Mountain View, CA 94040 (US). <u>GUEGLER</u> , Karl, J. [CH/US]; 1048 Oakland Avenue, Menlo Park, CA 94025 (US). <u>BAUGHN</u> , Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). <u>LU</u> , Dyung, Aina, M. [US/US]; 55 Park Belmont Place, San Jose, CA 95136 (US). <u>AZIMZAI</u> , Yalda [US/US]; 2045 Rock Springs Drive, Hayward, CA 94545 (US). <u>YANG</u> , Junming [CN/US]; 7136 Clarendon Street, San Jose, CA 95129 (US). (74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US). (81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published With international search report. (88) Date of publication of the international search report: 26 October 2000 (26.10.00)	
(22) International Filing Date: 12 November 1999 (12.11.99)			
(30) Priority Data:			
60/172,256	12 November 1998 (12.11.98) US		
60/135,519	21 May 1999 (21.05.99) US		
(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications			
US	60/172,256 (CIP)		
Filed on	12 November 1998 (12.11.98)		
US	60/135,519 (CIP)		
Filed on	21 May 1999 (21.05.99)		
(71) Applicant (for all designated States except US): INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).			
(72) Inventors; and			
(75) Inventors/Applicants (for US only): <u>TANG</u> , Tom, Y. [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). <u>HILLMAN</u> , Jennifer, L. [US/US]; 230 Monroe Drive, #12, Mountain View, CA 94040 (US). <u>YUE</u> , Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). <u>LAL</u> , Preeti [IN/US];			
(54) Title: HUMAN HYDROLASE PROTEINS			
(57) Abstract			
The invention provides human hydrolase proteins (HYDRL) and polynucleotides which identify and encode HYDRL. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of HYDRL.			

1	M	K	A	W	G	T	V	V	V	T	L	A	T	L	M	V	V	T	V	D	A	K	I	Y	E	L	C	E	L	A	R	L	E	R	2293764	
1	M	K	A	L	-	-	I	I	L	G	L	V	-	L	L	S	V	T	V	Q	G	K	I	F	E	R	C	E	L	A	R	T	L	K	K	g1790927
1	M	K	A	L	-	-	I	I	L	G	L	V	-	L	L	S	V	T	V	Q	G	K	I	F	E	R	C	E	L	A	R	T	L	K	K	g1790967
1	M	K	A	L	-	-	I	I	L	G	L	V	-	L	L	S	V	T	V	Q	G	K	I	F	E	R	C	E	L	A	R	T	L	K	K	g1790984
36	A	G	L	N	G	Y	K	G	Y	G	V	G	D	W	L	C	M	A	H	Y	E	S	G	F	D	T	A	F	V	D	H	N	P	-	D	2293764
33	L	G	L	D	G	Y	K	G	V	S	L	A	N	W	V	C	L	A	K	W	E	S	G	Y	N	T	D	A	T	N	Y	N	P	G	D	g1790927
33	L	G	L	D	G	Y	K	G	V	S	L	A	N	W	V	C	L	A	K	W	E	S	G	Y	N	T	D	A	T	N	Y	N	P	G	D	g1790967
33	L	G	L	D	G	Y	K	G	V	S	L	A	N	W	V	C	L	A	K	W	E	S	G	Y	N	T	E	A	T	N	Y	N	P	G	D	g1790984
70	G	S	S	E	Y	G	I	F	Q	L	N	S	A	W	C	D	N	G	I	T	P	T	K	-	N	L	C	H	M	D	C	H	D	L	2293764	
68	E	S	T	D	Y	G	I	F	Q	I	N	S	R	Y	W	C	N	N	G	K	T	P	G	A	V	N	A	C	H	I	S	C	N	A	L	g1790927
68	E	S	T	D	Y	G	I	F	Q	I	N	S	R	Y	W	C	N	N	G	K	T	P	G	A	V	N	A	C	H	I	S	C	N	A	L	g1790967
68	E	S	T	D	Y	G	I	F	Q	I	N	S	R	Y	W	C	N	N	G	K	T	P	G	A	V	N	A	C	H	I	S	C	S	A	L	g1790984
104	L	N	R	H	I	L	D	D	I	R	C	A	K	Q	I	V	S	S	Q	N	G	L	S	A	W	T	S	W	R	L	H	C	S	G	H	2293764
103	L	Q	N	N	I	A	D	A	V	A	C	A	K	R	V	V	S	D	P	Q	G	I	R	A	W	V	A	W	K	K	H	C	Q	N	R	g1790927
103	L	Q	N	N	I	A	D	A	V	A	C	A	K	R	V	V	S	D	P	Q	G	I	R	A	W	V	A	W	K	K	H	C	Q	N	R	g1790967
103	L	Q	N	N	I	A	D	A	V	A	C	A	K	R	V	V	S	D	P	Q	G	I	R	A	W	V	A	W	K	K	H	C	Q	N	R	g1790984
139	D	L	S	E	W	L	K	G	C	D	M	H	V	K	I	D	P	K	I	H	P															2293764
138	D	V	S	Q	Y	V	E	G	C	G	V																								g1790927	
138	D	V	S	Q	Y	V	E	G	C	G	V																								g1790967	
138	D	V	S	Q	Y	V	K	G	C	G	V																								g1790984	

FIGURE 1

Docket No.: PF-0634 USN

DECLARATION AND POWER OF ATTORNEY FOR UNITED STATES PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,
and

I believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if more than one name is listed below) of the subject matter which is claimed and for which a United States patent is sought on the invention entitled

HUMAN HYDROLASE PROTEINS

the specification of which:

 / is attached hereto.

 / was filed on _____ as application Serial No. _____ and if this box contains an X /, was amended on _____.

 / was filed as Patent Cooperation Treaty international application No. PCT/US99/27009 on November 12, 1999, if this box contains an X /, was amended on under Patent Cooperation Treaty Article 19 on _____ 2001, and if this box contains an X /, was amended on _____.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge my duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim the benefit under Title 35, United States Code, §119 or §365(a)-(b) of any foreign application(s) for patent or inventor's certificate indicated below and of any Patent Cooperation Treaty international applications(s) designating at least one country other than the United States indicated below and have also identified below any foreign application(s) for patent or inventor's certificate and Patent Cooperation Treaty international application(s) designating at least one country other than the United States for the same subject matter and having a filing date before that of the application for said subject matter the priority of which is claimed:

Country	Number	Filing Date	Priority Claimed
_____	_____	_____	<u> </u> / Yes <u> </u> / No
_____	_____	_____	<u> </u> / Yes <u> </u> / No

Docket No.: PF-0634 USN

I hereby claim the benefit under Title 35, United States Code, §119(e) of any United States provisional application(s) listed below.

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)
60/172,256	November 12, 1998	Expired
60/135,519	May 21, 1999	Expired

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in said prior application(s) in the manner required by the first paragraph of Title 35, United States Code §112, I acknowledge my duty to disclose material information as defined in Title 37 Code of Federal Regulations, §1.56(a) which occurred between the filing date(s) of the prior application(s) and the national or Patent Cooperation Treaty international filing date of this application:

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)
---------------------------	-------	--

I hereby appoint the following:

Lucy J. Billings	Reg. No. <u>36,749</u>
Michael C. Cerrone	Reg. No. <u>39,132</u>
Diana Hamlet-Cox	Reg. No. <u>33,302</u>
Richard C. Ekstrom	Reg. No. <u>37,027</u>
Barrie D. Greene	Reg. No. <u>46,740</u>
Matthew R. Kaser	Reg. No. <u>44,817</u>
Lynn E. Murry	Reg. No. <u>42,918</u>
Shirley A. Recipon	Reg. No. <u>47,016</u>
Susan K. Sather	Reg. No. <u>44,316</u>
Michelle M. Stempien	Reg. No. <u>41,327</u>
David G. Streeter	Reg. No. <u>43,168</u>
Stephen Todd	Reg. No. <u>47,139</u>
Christopher Turner	Reg. No. <u>45,167</u>
P. Ben Wang	Reg. No. <u>41,420</u>

respectively and individually, as my patent attorneys and/or agents, with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith. Please address all communications to:

Docket No.: PF-0634 USN

LEGAL DEPARTMENT
INCYTE GENOMICS, INC.
3160 PORTER DRIVE, PALO ALTO, CA 94304

TEL: 650-855-0555

FAX: 650-849-8886 or 650-845-4166

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Sole Inventor or

First Joint Inventor:

1-00

Full name:

Y. Tom Tang

Signature:

Y. Tom Tang

Date:

February 27, 2001

Citizenship

~~People's Republic of China~~ USA

Residence:

San Jose, California CA

P.O. Address:

4230 Ranwick Court
San Jose, California 95118

Second Joint Inventor:

2-00

Full name:

Jennifer L. Hillman

Signature:

Jennifer L. Hillman

Date:

February 16, 2001

Citizenship

United States of America

Residence:

Mountain View, California CA

P.O. Address:

230 Monroe Drive, #17
Mountain View, California 94040

Docket No.: PF-0634 USN

3-00

ThirdFourth Joint Inventor:

Full name:

Henry Yue/

Signature:

Henry Yue

Date:

March 07, 2001

Citizenship

United States of America

Residence:

Sunnyvale, California CA

P.O. Address:

826 Lois Avenue
Sunnyvale, California 94087

4-00

Fourth Joint Inventor:

Full name:

Preeti Lal

Signature:

Preeti Lal

Date:

FEBRUARY, 16, 2001

Citizenship

India

Residence:

Santa Clara, California CA

P.O. Address:

P.O. Box 5142
Santa Clara, California 95056

5-00

Fifth Joint Inventor:

Full name:

Olga Bandman

Signature:

Olga Bandman

Date:

16 February, 2001

Citizenship

United States of America

Residence:

Mountain View, California CA

P.O. Address:

366 Anna Avenue
Mountain View, California 94043

Docket No.: PF-0634 USN

6-00

Sixth Joint Inventor:

Full name:

Neil C. Corley

Signature:

Neil C. Corley

Date:

MARCH 5, 2001

Citizenship

United States of America

Residence:

Castro Valley, California CA

P.O. Address:

20426 Crow Creek Road
Castro Valley, California 94552

7-00

Seventh Joint Inventor:

Full name:

Karl J. Guegler

Signature:

K. J. Guegler

Date:

02/02, 2001

Citizenship

Switzerland

Residence:

Menlo Park, California CA

P.O. Address:

1048 Oakland Avenue
Menlo Park, California 94025

8-00

Eighth Joint Inventor:

Full name:

Mariah R. Baughn

Signature:

Mariah R. Baughn

Date:

February 12, 2001

Citizenship

United States of America

Residence:

San Leandro, California CA

P.O. Address:

14244 Santiago Road
San Leandro, California 94577

Docket No.: PF-0634 USN

9-00

Ninth Joint Inventor:

Full name:

Dyung Aina M. Lu

Signature:



Date:

March 22, 2001

Citizenship

United States of America

Residence:

San Jose, California CA

P.O. Address:

233 Coy Drive
San Jose, California 95123

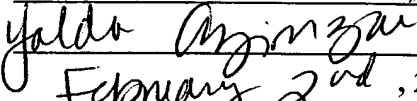
10-00

Tenth Joint Inventor:

Full name:

Yalda Azimzai

Signature:



Date:

February 2nd, 2001

Citizenship

United States of America

Residence:

Castro Valley, California CA

P.O. Address:

5518 Boulder Canyon Dr.
Castro Valley, California 94552

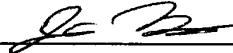
11-00

Eleventh Inventor:

Full name:

Junming Yang

Signature:



Date:

February 16, 2001

Citizenship

China

Residence:

Castro Valley, California CA

P.O. Address:

7125 Bark Lane
San Jose, California 95129

WO 00/28045

PCT/US99/27009

SEQUENCE LISTING

<110> INCYTE PHARMACEUTICALS, INC.

TANG, Y. Tom
 HILLMAN, Jennifer L.
 YUE, Henry
 LAL, Preeti
 BANDMAN, Olga
 CORLEY, Neil C.
 GUEGLER, Karl J.
 BAUGHN, Mariah R.
 LU, Dyung Aina M.
 AZIMZAI, Yalda
 YANG, Junming

<120> HUMAN HYDROLASE PROTEINS

<130> PF-0634 PCT

<140> To Be Assigned

<141> Herewith

<150> 09/190,937; unassigned; 60/135,519

<151> 1998-11-12; 1998-11-12; 1999-05-21

<160> 35

<170> PERL Program

<210> 1

<211> 159

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2293764CD1

<400> 1

Met	Lys	Ala	Trp	Gly	Thr	Val	Val	Val	Thr	Leu	Ala	Thr	Leu	Met
1				5					10					15
Val	Val	Thr	Val	Asp	Ala	Lys	Ile	Tyr	Glu	Leu	Cys	Glu	Leu	Ala
				20					25					30
Ala	Arg	Leu	Glu	Arg	Ala	Gly	Leu	Asn	Gly	Tyr	Lys	Gly	Tyr	Gly
				35					40					45
Val	Gly	Asp	Trp	Leu	Cys	Met	Ala	His	Tyr	Glu	Ser	Gly	Phe	Asp
				50					55					60
Thr	Ala	Phe	Val	Asp	His	Asn	Pro	Asp	Gly	Ser	Ser	Glu	Tyr	Gly
				65					70					75
Ile	Phe	Gln	Leu	Asn	Ser	Ala	Trp	Trp	Cys	Asp	Asn	Gly	Ile	Thr
				80					85					90
Pro	Thr	Lys	Asn	Leu	Cys	His	Met	Asp	Cys	His	Asp	Leu	Leu	Asn
				95					100					105
Arg	His	Ile	Leu	Asp	Asp	Ile	Arg	Cys	Ala	Lys	Gln	Ile	Val	Ser

[illegible]

```
<210> 2
<211> 285
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 949738CD1
```

<400>	2														
Met	Gly	Thr	Pro	Gly	Glu	Gly	Leu	Gly	Arg	Cys	Ser	His	Ala	Leu	
1				5					10					15	
Ile	Arg	Gly	Val	Pro	Glu	Ser	Leu	Ala	Ser	Gly	Glu	Gly	Ala	Gly	
				20					25					30	
Ala	Gly	Leu	Pro	Ala	Leu	Asp	Leu	Ala	Lys	Ala	Gln	Arg	Glu	His	
				35					40					45	
Gly	Val	Leu	Gly	Gly	Lys	Leu	Arg	Gln	Arg	Leu	Gly	Leu	Gln	Leu	
				50					55					60	
Leu	Glu	Leu	Pro	Pro	Glu	Glu	Ser	Leu	Pro	Leu	Gly	Pro	Leu	Leu	
				65					70					75	
Gly	Asp	Thr	Ala	Val	Ile	Gln	Gly	Asp	Thr	Ala	Leu	Ile	Thr	Arg	
				80					85					90	
Pro	Trp	Ser	Pro	Ala	Arg	Arg	Pro	Glu	Val	Asp	Gly	Val	Arg	Lys	
				95					100					105	
Ala	Leu	Gln	Asp	Leu	Gly	Leu	Arg	Ile	Val	Glu	Ile	Gly	Asp	Glu	
				110					115					120	
Asn	Ala	Thr	Leu	Asp	Gly	Thr	Asp	Val	Leu	Phe	Thr	Gly	Arg	Glu	
				125					130					135	
Phe	Phe	Val	Gly	Leu	Ser	Lys	Trp	Thr	Asn	His	Arg	Gly	Ala	Glu	
				140					145					150	
Ile	Val	Ala	Asp	Thr	Phe	Arg	Asp	Phe	Ala	Val	Ser	Thr	Val	Pro	
				155					160					165	
Val	Ser	Gly	Pro	Ser	His	Leu	Arg	Gly	Leu	Cys	Gly	Met	Gly	Gly	
				170					175					180	
Pro	Arg	Thr	Val	Val	Ala	Gly	Ser	Ser	Asp	Ala	Ala	Gln	Lys	Ala	
				185					190					195	
Val	Arg	Ala	Met	Ala	Val	Leu	Thr	Asp	His	Pro	Tyr	Ala	Ser	Leu	
				200					205					210	
Thr	Leu	Pro	Asp	Asp	Ala	Ala	Ala	Asp	Cys	Leu	Phe	Leu	Arg	Pro	
				215					220					225	
Gly	Leu	Pro	Gly	Val	Pro	Pro	Phe	Leu	Leu	His	Arg	Gly	Gly	Gly	
				230					235					240	
Asp	Leu	Pro	Asn	Ser	Gln	Glu	Ala	Leu	Gln	Lys	Leu	Ser	Asp	Val	
				245					250					255	
Thr	Leu	Val	Pro	Val	Ser	Cys	Ser	Glu	Leu	Glu	Lys	Ala	Gly	Ala	
				260					265					270	

WO 00/28045

PCT/US99/27009

Gly Leu Ser Ser Leu Cys Leu Val Leu Ser Thr Arg Pro His Ser
 275 280 285

<210> 3

<211> 331

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1297034CD1

<400> 3

Met Trp Leu Trp Glu Asp Gln Gly Gly Leu Leu Gly Pro Phe Ser
 1 5 10 15
 Phe Leu Leu Leu Val Leu Leu Leu Val Thr Arg Ser Pro Val Asn
 20 25 30
 Ala Cys Leu Leu Thr Gly Ser Leu Phe Val Leu Leu Arg Val Phe
 35 40 45
 Ser Phe Glu Pro Val Pro Ser Cys Arg Ala Leu Gln Val Leu Lys
 50 55 60
 Pro Arg Asp Arg Ile Ser Ala Ile Ala His Arg Gly Gly Ser His
 65 70 75
 Asp Ala Pro Glu Asn Thr Leu Ala Ala Ile Arg Gln Ala Ala Lys
 80 85 90
 Asn Gly Ala Thr Gly Val Glu Leu Asp Ile Glu Phe Thr Ser Asp
 95 100 105
 Gly Ile Pro Val Leu Met His Asp Asn Thr Val Asp Arg Thr Thr
 110 115 120
 Asp Gly Thr Gly Arg Leu Cys Asp Leu Thr Phe Glu Gln Ile Arg
 125 130 135
 Lys Leu Asn Pro Ala Ala Asn His Arg Leu Arg Asn Asp Phe Pro
 140 145 150
 Asp Glu Lys Ile Pro Thr Leu Arg Glu Ala Val Ala Glu Cys Leu
 155 160 165
 Asn His Asn Leu Thr Ile Phe Phe Asp Val Lys Gly His Ala His
 170 175 180
 Lys Ala Thr Glu Ala Leu Lys Lys Met Tyr Met Glu Phe Pro Gln
 185 190 195
 Leu Tyr Asn Asn Ser Val Val Cys Ser Phe Leu Pro Glu Val Ile
 200 205 210
 Tyr Lys Met Arg Gln Thr Asp Arg Asp Val Ile Thr Ala Leu Thr
 215 220 225
 His Arg Pro Trp Ser Leu Ser His Thr Gly Asp Gly Lys Pro Arg
 230 235 240
 Tyr Asp Thr Phe Trp Lys His Phe Ile Phe Val Met Met Asp Ile
 245 250 255
 Leu Leu Asp Trp Ser Met His Asn Ile Leu Trp Tyr Leu Cys Gly
 260 265 270
 Ile Ser Ala Phe Leu Met Gln Lys Asp Phe Val Ser Pro Ala Tyr
 275 280 285
 Leu Lys Lys Trp Ser Ala Lys Gly Ile Gln Val Val Gly Trp Thr
 290 295 300
 Val Asn Thr Phe Asp Glu Lys Ser Tyr Tyr Glu Ser His Leu Gly

```

          305                      310                      315
Ser Ser Tyr Ile Thr Asp Ser Met Val Glu Asp Cys Glu Pro His
          320                      325                      330
Phe

```

```
<210> 4
<211> 153
<212> PRT
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<223> Incyte ID No: 1553276CD1
```

[illegible]

```
<210> 5
<211> 571
<212> PRT
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<223> Incyte ID No: 1702211CD1
```

```

<400> 5
Met Glu Arg Ala Val Arg Val Glu Ser Gly Val Leu Val Gly Val
  1                      5          10          15
Val Cys Leu Leu Leu Ala Cys Pro Ala Thr Ala Thr Gly Pro Glu
                20          25          30
Val Ala Gln Pro Glu Val Asp Thr Thr Leu Gly Arg Val Arg Gly
                35          40          45

```

Arg	Gln	Val	Gly	Val	Lys	Gly	Thr	Asp	Arg	Leu	Val	Asn	Val	Phe	60
				50					55						60
Leu	Gly	Ile	Pro	Phe	Ala	Gln	Pro	Pro	Leu	Gly	Pro	Asp	Arg	Phe	65
				65					70						75
Ser	Ala	Pro	His	Pro	Ala	Gln	Pro	Trp	Glu	Gly	Val	Arg	Asp	Ala	80
				80					85						90
Ser	Thr	Ala	Pro	Pro	Met	Cys	Leu	Gln	Asp	Val	Glu	Ser	Met	Asn	95
				95					100						105
Ser	Ser	Arg	Phe	Val	Leu	Asn	Gly	Lys	Gln	Gln	Ile	Phe	Ser	Val	110
				110					115						120
Ser	Glu	Asp	Cys	Leu	Val	Leu	Asn	Val	Tyr	Ser	Pro	Ala	Glu	Val	125
				125					130						135
Pro	Ala	Gly	Ser	Gly	Arg	Pro	Val	Met	Val	Trp	Val	His	Gly	Gly	140
				140					145						150
Ala	Leu	Ile	Thr	Gly	Ala	Ala	Thr	Ser	Tyr	Asp	Gly	Ser	Ala	Leu	155
				155					160						165
Ala	Ala	Tyr	Gly	Asp	Val	Val	Val	Val	Thr	Val	Gln	Tyr	Arg	Leu	170
				170					175						180
Gly	Val	Leu	Gly	Phe	Phe	Ser	Thr	Gly	Asp	Glu	His	Ala	Pro	Gly	185
				185					190						195
Asn	Gln	Gly	Phe	Leu	Asp	Val	Val	Ala	Ala	Leu	Arg	Trp	Val	Gln	200
				200					205						210
Glu	Asn	Ile	Ala	Pro	Phe	Gly	Gly	Asp	Leu	Asn	Cys	Val	Thr	Val	215
				215					220						225
Phe	Gly	Gly	Ser	Ala	Gly	Gly	Ser	Ile	Ile	Ser	Gly	Leu	Val	Leu	230
				230					235						240
Ser	Pro	Val	Ala	Ala	Gly	Leu	Phe	His	Arg	Ala	Ile	Thr	Gln	Ser	245
				245					250						255
Gly	Val	Ile	Thr	Thr	Pro	Gly	Ile	Ile	Asp	Ser	His	Pro	Trp	Pro	260
				260					265						270
Leu	Ala	Gln	Lys	Ile	Ala	Asn	Thr	Leu	Ala	Cys	Ser	Ser	Ser	Ser	275
				275					280						285
Pro	Ala	Glu	Met	Val	Gln	Cys	Leu	Gln	Gln	Lys	Glu	Gly	Glu	Glu	290
				290					295						300
Leu	Val	Leu	Ser	Lys	Lys	Leu	Lys	Asn	Thr	Ile	Tyr	Pro	Leu	Thr	305
				305					310						315
Val	Asp	Gly	Thr	Val	Phe	Pro	Lys	Ser	Pro	Lys	Glu	Leu	Leu	Lys	320
				320					325						330
Glu	Lys	Pro	Phe	His	Ser	Val	Pro	Phe	Leu	Met	Gly	Val	Asn	Asn	335
				335					340						345
His	Glu	Phe	Ser	Trp	Leu	Ile	Pro	Arg	Gly	Trp	Gly	Leu	Leu	Asp	350
				350					355						360
Thr	Met	Glu	Gln	Met	Ser	Arg	Glu	Asp	Met	Leu	Ala	Ile	Ser	Thr	365
				365					37						

WO 00/28045

PCT/US99/27009

	455	460	465
Phe Val Phe Gly Gly Pro Phe Leu Met Asp Glu Ser Ser Arg Leu			
	470	475	480
Ala Phe Pro Glu Ala Thr Glu Glu Glu Lys Gln Leu Ser Leu Thr			
	485	490	495
Met Met Ala Gln Trp Thr His Phe Ala Arg Thr Gly Asp Pro Asn			
	500	505	510
Ser Lys Ala Leu Pro Pro Trp Pro Gln Phe Asn Gln Ala Glu Gln			
	515	520	525
Tyr Leu Glu Ile Asn Pro Val Pro Arg Ala Gly Gln Lys Phe Arg			
	530	535	540
Glu Ala Trp Met Gln Phe Trp Ser Glu Thr Leu Pro Ser Lys Ile			
	545	550	555
Gln Gln Trp His Gln Lys Gln Lys Asn Arg Lys Ala Gln Glu Asp			
	560	565	570
Leu			

<210> 6

<211> 347

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1859618CD1

<400> 6

Met Ser Ser Trp Ser Arg Gln Arg Pro Lys Ser Pro Gly Gly Ile		
1	5	10
Gln Pro His Val Ser Arg Thr Leu Phe Leu Leu Leu Leu Ala		15
	20	25
Ala Ser Ala Trp Gly Val Thr Leu Ser Pro Lys Asp Cys Gln Val		30
	35	40
Phe Arg Ser Asp His Gly Ser Ser Ile Ser Cys Gln Pro Pro Ala		45
	50	55
Glu Ile Pro Gly Tyr Leu Pro Ala Asp Thr Val His Leu Ala Val		60
	65	70
Glu Phe Phe Asn Leu Thr His Leu Pro Ala Asn Leu Leu Gln Gly		75
	80	85
Ala Ser Lys Leu Gln Glu Leu His Leu Ser Ser Asn Gly Leu Glu		90
	95	100
Ser Leu Ser Pro Glu Phe Leu Arg Pro Val Pro Gln Leu Arg Val		105
	110	115
Leu Asp Leu Thr Arg Asn Ala Leu Thr Gly Leu Pro Pro Gly Leu		120
	125	130
Phe Gln Ala Ser Ala Thr Leu Asp Thr Leu Val Leu Lys Glu Asn		135
	140	145
Gln Leu Glu Val Leu Glu Val Ser Trp Leu His Gly Leu Lys Ala		150
	155	160
Leu Gly His Leu Asp Leu Ser Gly Asn Arg Leu Arg Lys Leu Pro		165
	170	175
Pro Gly Leu Leu Ala Asn Phe Thr Leu Leu Arg Thr Leu Asp Leu		180
	185	190
Gly Glu Asn Gln Leu Glu Thr Leu Pro Pro Asp Leu Leu Arg Gly		195

WO 00/28045

PCT/US99/27009

200	205	210
Pro Leu Gln Leu Glu Arg Leu His Leu Glu Gly Asn Lys Leu Gln		
215	220	225
Val Leu Gly Lys Asp Leu Leu Leu Pro Gln Pro Asp Leu Arg Tyr		
230	235	240
Leu Phe Leu Asn Gly Asn Lys Leu Ala Arg Val Ala Ala Gly Ala		
245	250	255
Phe Gln Gly Leu Arg Gln Leu Asp Met Leu Asp Leu Ser Asn Asn		
260	265	270
Ser Leu Ala Ser Val Pro Glu Gly Leu Trp Ala Ser Leu Gly Gln		
275	280	285
Pro Asn Trp Asp Met Arg Asp Gly Phe Asp Ile Ser Gly Asn Pro		
290	295	300
Trp Ile Cys Asp Gln Asn Leu Ser Asp Leu Tyr Arg Trp Leu Gln		
305	310	315
Ala Gln Lys Asp Lys Met Phe Ser Gln Asn Asp Thr Arg Cys Ala		
320	325	330
Gly Pro Glu Ala Val Lys Gly Gln Thr Leu Leu Ala Val Ala Lys		
335	340	345
Ser Gln		

<210> 7

<211> 194

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2011071CD1

<400> 7

Met Gln Asp Ala Pro Leu Ser Cys Leu Ser Pro Thr Arg Trp Ser		
1	5	10
Ser Val Ser Ser Ala Asp Ser Thr Glu Lys Ser Ala Ser Gly Ala		
20	25	30
Gly Thr Arg Asn Leu Pro Phe Gln Phe Cys Leu Arg Gln Ala Leu		
35	40	45
Arg Met Lys Ala Ala Gly Ile Leu Thr Leu Ile Gly Cys Leu Val		
50	55	60
Thr Gly Ala Glu Ser Lys Ile Tyr Thr Arg Cys Lys Leu Ala Lys		
65	70	75
Ile Phe Ser Arg Ala Gly Leu Asp Asn Tyr Trp Gly Phe Ser Leu		
80	85	90
Gly Asn Trp Ile Cys Met Ala Tyr Tyr Glu Ser Gly Tyr Asn Thr		
95	100	105
Thr Ala Pro Thr Val Leu Asp Asp Gly Ser Ile Asp Tyr Gly Ile		
110	115	120
Phe Gln Ile Asn Thr Phe Ala Trp Cys Arg Arg Gly Lys Leu Lys		
125	130	135
Glu Asn Asn His Cys His Val Ala Cys Ser Ala Leu Ile Thr Asp		
140	145	150
Asp Leu Thr Asp Ala Ile Ile Cys Ala Arg Lys Ile Val Lys Glu		
155	160	165
Thr Gln Gly Met Asn Tyr Trp Gln Gly Trp Lys Lys His Cys Glu		

WO 00/28045

PCT/US99/27009

170 175 180
 Gly Arg Asp Leu Ser Glu Trp Lys Lys Gly Cys Glu Val Ser
 185 190

<210> 8

<211> 361

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2186517CD1

<400> 8

Met Ala Trp Gln Gly Trp Pro Ala Ala Trp Gln Trp Val Ala Gly
 1 5 10 15
 Cys Trp Leu Leu Leu Val Leu Val Leu Val Leu Leu Val Ser Pro
 20 25 30
 Arg Gly Cys Arg Ala Arg Arg Gly Leu Arg Gly Leu Leu Met Ala
 35 40 45
 His Ser Gln Arg Leu Leu Phe Arg Ile Gly Tyr Ser Leu Tyr Thr
 50 55 60
 Arg Thr Trp Leu Gly Tyr Leu Phe Tyr Arg Gln Gln Leu Arg Arg
 65 70 75
 Ala Arg Asn Arg Tyr Pro Lys Gly His Ser Lys Thr Gln Thr Arg
 80 85 90
 Leu Phe Asn Gly Val Lys Val Leu Pro Ile Pro Val Leu Ser Asp
 95 100 105
 Asn Tyr Ser Tyr Leu Ile Ile Asp Thr Gln Ala Gln Leu Ala Val
 110 115 120
 Ala Val Asp Pro Ser Asp Pro Arg Ala Val Gln Ala Ser Ile Glu
 125 130 135
 Lys Glu Gly Val Thr Leu Val Ala Ile Leu Cys Thr His Lys His
 140 145 150
 Trp Asp His Ser Gly Gly Asn Arg Asp Leu Ser Arg Arg His Arg
 155 160 165
 Asp Cys Arg Val Tyr Gly Ser Pro Gln Asp Gly Ile Pro Tyr Leu
 170 175 180
 Thr His Pro Leu Cys His Gln Asp Val Val Ser Val Gly Arg Leu
 185 190 195
 Gln Ile Arg Ala Leu Ala Thr Pro Gly His Thr Gln Gly His Leu
 200 205 210
 Val Tyr Leu Leu Asp Gly Glu Pro Tyr Lys Gly Pro Ser Cys Leu
 215 220 225
 Phe Ser Gly Asp Leu Leu Phe Leu Ser Gly Cys Gly Arg Thr Phe
 230 235 240
 Glu Gly Asn Ala Glu Thr Met Leu Ser Ser Leu Asp Thr Val Leu
 245 250 255
 Gly Leu Gly Asp Asp Thr Leu Leu Trp Pro Gly His Glu Tyr Ala
 260 265 270
 Glu Glu Asn Leu Gly Phe Ala Gly Val Val Glu Pro Glu Asn Leu
 275 280 285
 Ala Arg Glu Arg Lys Met Gln Trp Val Gln Arg Gln Arg Leu Glu
 290 295 300

WO 00/28045

PCT/US99/27009

Arg	Lys	Gly	Thr	Cys	Pro	Ser	Thr	Leu	Gly	Glu	Glu	Arg	Ser	Tyr
				305					310					315
Asn	Pro	Phe	Leu	Arg	Thr	His	Cys	Leu	Ala	Leu	Gln	Glu	Ala	Leu
				320					325					330
Gly	Pro	Gly	Pro	Gly	Pro	Thr	Gly	Asp	Asp	Asp	Tyr	Ser	Arg	Ala
				335					340					345
Gln	Leu	Leu	Glu	Glu	Leu	Arg	Arg	Leu	Lys	Asp	Met	His	Lys	Ser
				350					355					360

Lys

<210> 9

<211> 306

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2253585CD1

<400> 9

Met	Leu	Arg	Trp	Thr	Arg	Ala	Trp	Arg	Leu	Pro	Arg	Glu	Gly	Leu
1				5					10					15
Gly	Pro	His	Gly	Pro	Ser	Phe	Ala	Arg	Val	Pro	Val	Ala	Pro	Ser
				20					25					30
Ser	Ser	Ser	Gly	Gly	Arg	Gly	Gly	Ala	Glu	Pro	Arg	Pro	Leu	Pro
				35					40					45
Leu	Ser	Tyr	Arg	Leu	Leu	Asp	Gly	Glu	Ala	Ala	Leu	Pro	Ala	Val
				50					55					60
Val	Phe	Leu	His	Gly	Leu	Phe	Gly	Ser	Lys	Thr	Asn	Phe	Asn	Ser
				65					70					75
Ile	Ala	Lys	Ile	Leu	Ala	Gln	Gln	Thr	Gly	Arg	Arg	Val	Leu	Thr
				80					85					90
Val	Asp	Ala	Arg	Asn	His	Gly	Asp	Ser	Pro	His	Ser	Pro	Asp	Met
				95					100					105
Ser	Tyr	Glu	Ile	Met	Ser	Gln	Asp	Leu	Gln	Asp	Leu	Leu	Pro	Gln
				110					115					120
Leu	Gly	Leu	Val	Pro	Cys	Val	Val	Val	Gly	His	Ser	Met	Gly	Gly
				125					130					135
Lys	Thr	Ala	Met	Leu	Leu	Ala	Leu	Gln	Arg	Pro	Glu	Leu	Val	Glu
				140					145					150
Arg	Leu	Ile	Ala	Val	Asp	Ile	Ser	Pro	Val	Glu	Ser	Thr	Gly	Val
				155					160					165
Ser	His	Phe	Ala	Thr	Tyr	Val	Ala	Ala	Met	Arg	Ala	Ile	Asn	Ile
				170					175					180
Ala	Asp	Glu	Leu	Pro	Arg	Ser	Arg	Ala	Arg	Lys	Leu	Ala	Asp	Glu
				185					190					195
Gln	Leu	Ser	Ser	Val	Ile	Gln	Asp	Met	Ala	Val	Arg	Gln	His	Leu
				200					205					210
Leu	Thr	Asn	Leu	Val	Glu	Val	Asp	Gly	Arg	Phe	Val	Trp	Arg	Val
				215					220					225
Asn	Leu	Asp	Ala	Leu	Thr	Gln	His	Leu	Asp	Lys	Ile	Leu	Ala	Phe
				230					235					240
Pro	Gln	Arg	Gln	Glu	Ser	Tyr	Leu	Gly	Pro	Thr	Leu	Phe	Leu	Leu
				245					250					255

WO 00/28045

PCT/US99/27009

Gly Gly Asn Ser Gln Phe Val His Pro Ser His His Pro Glu Ile		
	260	265 270
Met Arg Leu Phe Pro Arg Ala Gln Met Gln Thr Val Pro Asn Ala		
	275	280 285
Gly His Trp Ile His Ala Asp Arg Pro Gln Asp Phe Ile Ala Ala		
	290	295 300
Ile Arg Gly Phe Leu Val		
	305	

<210> 10
 <211> 483
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2447520CD1

<400> 10

Met Ser Asn Lys Leu Leu Ser Pro His Pro His Ser Val Val Leu		
1	5	10 15
Arg Ser Glu Phe Lys Met Ala Ser Ser Pro Ala Val Leu Arg Ala		
	20	25 30
Ser Arg Leu Tyr Gln Trp Ser Leu Lys Ser Ser Ala Gln Phe Leu		
	35	40 45
Gly Ser Pro Gln Leu Arg Gln Val Gly Gln Ile Ile Arg Val Pro		
	50	55 60
Ala Arg Met Ala Ala Thr Leu Ile Leu Glu Pro Ala Gly Arg Cys		
	65	70 75
Cys Trp Asp Glu Pro Val Arg Ile Ala Val Arg Gly Leu Ala Pro		
	80	85 90
Glu Gln Pro Val Thr Leu Arg Ala Ser Leu Arg Asp Glu Lys Gly		
	95	100 105
Ala Leu Phe Gln Ala His Ala Arg Tyr Arg Ala Asp Thr Leu Gly		
	110	115 120
Glu Leu Asp Leu Glu Arg Ala Pro Ala Leu Gly Gly Ser Phe Ala		
	125	130 135
Gly Leu Glu Pro Met Gly Leu Leu Trp Ala Leu Glu Pro Glu Lys		
	140	145 150
Pro Leu Val Arg Leu Val Lys Arg Asp Val Arg Thr Pro Leu Ala		
	155	160 165
Val Glu Leu Glu Val Leu Asp Gly His Asp Pro Asp Pro Gly Arg		
	170	175 180
Leu Leu Cys Gln Thr Arg His Glu Arg Tyr Phe Leu Pro Pro Gly		
	185	190 195
Val Arg Arg Glu Pro Val Arg Val Gly Arg Val Arg Gly Thr Leu		
	200	205 210
Phe Leu Pro Pro Glu Pro Gly Pro Phe Pro Gly Ile Val Asp Met		
	215	220 225
Phe Gly Thr Gly Gly Gly Leu Leu Glu Tyr Arg Ala Ser Leu Leu		
	230	235 240
Ala Gly Lys Gly Phe Ala Val Met Ala Leu Ala Tyr Tyr Asn Tyr		
	245	250 255
Glu Asp Leu Pro Lys Thr Met Glu Thr Leu His Leu Glu Tyr Phe		

WO 00/28045

PCT/US99/27009

	260		265		270
Glu Glu Ala Met	Asn Tyr Leu Leu Ser	His Pro Glu Val Lys Gly			
	275		280		285
Pro Gly Val Gly	Leu Leu Gly Ile Ser	Lys Gly Gly Glu Leu Cys			
	290		295		300
Leu Ser Met Ala	Ser Phe Leu Lys Gly	Ile Thr Ala Ala Val Val			
	305		310		315
Ile Asn Gly Ser	Val Ala Asn Val Gly	Gly Thr Leu Arg Tyr Lys			
	320		325		330
Gly Glu Thr Leu	Pro Pro Val Gly Val	Asn Arg Asn Arg Ile Lys			
	335		340		345
Val Thr Lys Asp	Gly Tyr Ala Asp Ile	Val Asp Val Leu Asn Ser			
	350		355		360
Pro Leu Glu Gly	Pro Asp Gln Lys Ser	Phe Ile Pro Val Glu Arg			
	365		370		375
Ala Glu Ser Thr	Phe Leu Phe Leu Val	Gly Gln Asp Asp His Asn			
	380		385		390
Trp Lys Ser Glu	Phe Tyr Ala Asn Glu	Ala Cys Lys Arg Leu Gln			
	395		400		405
Ala His Gly Arg	Arg Lys Pro Gln Ile	Ile Cys Tyr Pro Glu Thr			
	410		415		420
Gly His Tyr Ile	Glu Pro Pro Tyr Phe	Pro Leu Cys Arg Ala Ser			
	425		430		435
Leu His Ala Leu	Val Gly Ser Pro Ile	Ile Trp Gly Gly Glu Pro			
	440		445		450
Arg Ala His Ala	Met Ala Gln Val Asp	Ala Trp Lys Gln Leu Gln			
	455		460		465
Thr Phe Phe His	Lys His Leu Gly Gly	His Glu Gly Thr Ile Pro			
	470		475		480
Ser Lys Val					

<210> 11
 <211> 144
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2481345CD1

<400> 11
 Met Leu Leu Leu Trp Val Ser Val Val Ala Ala Leu Ala Leu Ala
 1 5 10 15
 Val Leu Ala Pro Gly Ala Gly Glu Gln Arg Arg Arg Ala Ala Lys
 20 25 30
 Ala Pro Asn Val Val Leu Val Val Ser Asp Ser Phe Asp Gly Arg
 35 40 45
 Leu Thr Phe His Pro Gly Ser Gln Val Val Lys Leu Pro Phe Ile
 50 55 60
 Asn Phe Met Lys Thr Arg Gly Thr Ser Phe Leu Asn Ala Tyr Thr
 65 70 75
 Asn Ser Pro Ile Cys Cys Pro Ser Arg Ala Ala Met Trp Ser Gly
 80 85 90
 Leu Phe Thr His Leu Thr Glu Ser Trp Asn Asn Phe Lys Gly Leu

[illegible]

```
<210> 12
<211> 180
<212> PRT
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<223> Incyte ID No: 2484020CD1
```

<400> 12															
Met	Met	Lys	Phe	Lys	Pro	Asn	Gln	Thr	Arg	Thr	Tyr	Asp	Arg	Glu	
1				5					10					15	
Gly	Phe	Lys	Lys	Arg	Ala	Ala	Cys	Leu	Cys	Phe	Arg	Ser	Glu	Gln	
				20					25					30	
Glu	Asp	Glu	Val	Leu	Leu	Val	Ser	Ser	Ser	Arg	Tyr	Pro	Asp	Gln	
				35					40					45	
Trp	Ile	Val	Pro	Gly	Gly	Gly	Met	Glu	Pro	Glu	Glu	Glu	Pro	Gly	
				50					55					60	
Gly	Ala	Ala	Val	Arg	Glu	Val	Tyr	Glu	Glu	Ala	Gly	Val	Lys	Gly	
				65					70					75	
Lys	Leu	Gly	Arg	Leu	Leu	Gly	Ile	Phe	Glu	Asn	Gln	Asp	Arg	Lys	
				80					85					90	
His	Arg	Thr	Tyr	Val	Tyr	Val	Leu	Thr	Val	Thr	Glu	Ile	Leu	Glu	
				95					100					105	
Asp	Trp	Glu	Asp	Ser	Val	Asn	Ile	Gly	Arg	Lys	Arg	Glu	Trp	Phe	
				110					115					120	
Lys	Val	Glu	Asp	Ala	Ile	Lys	Val	Leu	Gln	Cys	His	Lys	Pro	Val	
				125					130					135	
His	Ala	Glu	Tyr	Leu	Glu	Lys	Leu	Lys	Leu	Gly	Cys	Ser	Pro	Ala	
				140					145					150	
Asn	Gly	Asn	Ser	Thr	Val	Pro	Ser	Leu	Pro	Asp	Asn	Asn	Ala	Leu	
				155					160					165	
Phe	Val	Thr	Ala	Ala	Gln	Thr	Ser	Gly	Leu	Pro	Ser	Ser	Val	Arg	
				170					175					180	

```
<210> 13
<211> 375
<212> PRT
<213> Homo sapiens
```

```
<220>
<221> misc_feature
<223> Incyte ID No: 2862528CD1
```


<212> PRT

<213> Homo sapiens

<220>

```
<221> misc_feature
```

<223> Incyte ID No: 3200650CD1

<400> 14

Met	Thr	Thr	Trp	Ser	Leu	Arg	Arg	Arg	Pro	Ala	Arg	Thr	Leu	Gly
1				5					10					15
Leu	Leu	Leu	Leu	Val	Val	Leu	Gly	Phe	Leu	Val	Leu	Arg	Arg	Leu
				20					25					30
Asp	Trp	Ser	Thr	Leu	Val	Pro	Leu	Arg	Leu	Arg	His	Arg	Gln	Leu
				35					40					45
Gly	Leu	Gln	Ala	Lys	Gly	Trp	Asn	Phe	Met	Leu	Glu	Asp	Ser	Thr
				50					55					60
Phe	Trp	Ile	Phe	Gly	Gly	Ser	Ile	His	Tyr	Phe	Arg	Val	Pro	Arg
				65					70					75
Glu	Tyr	Trp	Arg	Asp	Arg	Leu	Leu	Lys	Met	Lys	Ala	Cys	Gly	Leu
				80					85					90
Asn	Thr	Leu	Thr	Thr	Tyr	Val	Pro	Trp	Asn	Leu	His	Glu	Pro	Glu
				95					100					105
Arg	Gly	Lys	Phe	Asp	Phe	Leu	Trp	Glu	Thr	Trp	Thr	Leu	Lys	Ala
				110					115					120
Phe	Val	Leu	Met	Ala	Ala	Glu	Ile	Gly	Leu	Trp	Val	Ile	Leu	Arg
				125					130					135
Pro	Gly	Pro	Tyr	Ile	Cys	Ser	Glu	Met	Asp	Leu	Gly	Gly	Leu	Pro
				140					145					150
Ser	Trp	Leu	Leu	Gln	Asp	Pro	Gly	Met	Arg	Leu	Arg	Thr	Thr	Tyr
				155					160					165
Lys	Gly	Phe	Thr	Glu	Ala	Val	Asp	Leu	Tyr	Phe	Asp	His	Leu	Met
				170					175					180
Ser	Arg	Val	Val	Pro	Leu	Gln	Tyr	Lys	Arg	Gly	Gly	Pro	Ile	Ile
				185					190					195
Ala	Val	Gln	Val	Glu	Asn	Glu	Tyr	Gly	Ser	Tyr	Asn	Lys	Asp	Pro
				200					205					210
Ala	Tyr	Met	Pro	Tyr	Val	Lys	Lys	Ala	Leu	Glu	Asp	Arg	Gly	Ile
				215					220					225
Val	Glu	Leu	Leu	Leu	Thr	Ser	Asp	Asn	Lys	Asp	Gly	Leu	Ser	Lys
				230					235					240
Gly	Ile	Val	Gln	Gly	Val	Leu	Ala	Thr	Ile	Asn	Leu	Gln	Ser	Thr
				245					250					255
His	Glu	Leu	Gln	Leu	Leu	Thr	Thr	Phe	Leu	Phe	Asn	Val	Gln	Gly
				260					265					270
Thr	Gln	Pro	Lys	Met	Val	Met	Glu	Tyr	Trp	Thr	Gly	Trp	Phe	Asp
				275					280					285
Ser	Trp	Gly	Gly	Pro	His	Asn	Ile	Leu	Asp	Ser	Ser	Glu	Val	Leu
				290					295					300
Lys	Thr	Val	Ser	Ala	Ile	Val	Asp	Ala	Gly	Ser	Ser	Ile	Asn	Leu
				305					310					315
Tyr	Met	Phe	His	Gly	Gly	Thr	Asn	Phe	Gly	Phe	Met	Asn	Gly	Ala
				320					325					330
Met	His	Phe	His	Asp	Tyr	Lys	Ser	Asp	Val	Thr	Ser	Tyr	Asp	Tyr
				335					340					345
Asp	Ala	Val	Leu	Thr	Glu	Ala	Gly	Asp	Tyr	Thr	Ala	Lys	Tyr	Met

15/30

WO 00/28045

PCT/US99/27009

Gln Gly Pro Pro Val Leu Cys Leu His Gly Trp Leu Asp Asn Ala
 35 40 45
 Ser Ser Phe Asp Arg Leu Ile Pro Leu Leu Pro Gln Asp Phe Tyr
 50 55 60
 Tyr Val Ala Met Asp Phe Gly Gly His Gly Leu Ser Ser His Tyr
 65 70 75
 Ser Pro Gly Val Pro Tyr Tyr Leu Gln Thr Phe Val Ser Glu Ile
 80 85 90
 Arg Arg Val Val Ala Ala Leu Lys Trp Asn Arg Phe Ser Ile Leu
 95 100 105
 Gly His Ser Phe Gly Gly Val Val Gly Gly Met Phe Phe Cys Thr
 110 115 120
 Phe Pro Glu Met Val Asp Lys Leu Ile Leu Leu Asp Thr Pro Leu
 125 130 135
 Phe Leu Leu Glu Ser Asp Glu Met Glu Asn Leu Leu Thr Tyr Lys
 140 145 150
 Arg Arg Ala Ile Glu His Val Leu Gln Val Glu Ala Ser Gln Glu
 155 160 165
 Pro Ser His Val Phe Ser Leu Lys Gln Leu Leu Gln Arg Leu Leu
 170 175 180
 Lys Ser Asn Ser His Leu Ser Glu Glu Cys Gly Glu Leu Leu Leu
 185 190 195
 Gln Arg Gly Thr Thr Lys Val Ala Thr Gly Leu Val Leu Asn Arg
 200 205 210
 Asp Gln Arg Leu Ala Trp Ala Glu Asn Ser Ile Asp Phe Ile Ser
 215 220 225
 Arg Glu Leu Cys Ala His Ser Ile Arg Lys Leu Gln Ala His Val
 230 235 240
 Leu Leu Ile Lys Ala Val His Gly Tyr Phe Asp Ser Arg Gln Asn
 245 250 255
 Tyr Ser Glu Lys Glu Ser Leu Ser Phe Met Ile Asp Thr Met Lys
 260 265 270
 Ser Thr Leu Lys Glu Gln Phe Gln Phe Val Glu Val Pro Gly Asn
 275 280 285
 His Cys Val His Met Ser Glu Pro Gln His Val Ala Ser Ile Ile
 290 295 300
 Ser Ser Phe Leu Gln Cys Thr His Met Leu Pro Ala Gln Leu
 305 310

<210> 16

<211> 448

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4661133CD1

<400> 16

Met Arg Arg Ala Ala Leu Arg Leu Cys Ala Leu Gly Lys Gly Gln
 1 5 10 15
 Leu Thr Pro Gly Arg Gly Leu Thr Gln Gly Pro Gln Asn Pro Lys
 20 25 30
 Lys Gln Gly Ile Phe His Ile His Glu Ala Cys Ser Ser Ile His

					35					40					45
Val	Asn	His	Val	Arg	Asp	Lys	Leu	Arg	Glu	Ile	Val	Gly	Ala	Ser	
					50					55					60
Thr	Asn	Trp	Arg	Asp	His	Val	Lys	Ala	Met	Glu	Glu	Arg	Lys	Leu	
					65					70					75
Leu	His	Ser	Phe	Leu	Ala	Lys	Ser	Gln	Asp	Gly	Leu	Pro	Pro	Arg	
					80					85					90
Arg	Met	Lys	Asp	Ser	Tyr	Ile	Glu	Val	Leu	Leu	Pro	Leu	Gly	Ser	
					95					100					105
Glu	Pro	Glu	Leu	Arg	Glu	Lys	Tyr	Leu	Thr	Val	Gln	Asn	Thr	Val	
					110					115					120
Arg	Phe	Gly	Arg	Ile	Leu	Glu	Asp	Leu	Asp	Ser	Leu	Gly	Val	Leu	
					125					130					135
Ile	Cys	Tyr	Met	His	Asn	Lys	Ile	His	Ser	Ala	Lys	Met	Ser	Pro	
					140					145					150
Leu	Ser	Ile	Val	Thr	Ala	Leu	Val	Asp	Lys	Ile	Asp	Met	Cys	Lys	
					155					160					165
Lys	Ser	Leu	Ser	Pro	Glu	Gln	Asp	Ile	Lys	Phe	Ser	Gly	His	Val	
					170					175					180
Ser	Trp	Val	Gly	Lys	Thr	Ser	Met	Glu	Val	Lys	Met	Gln	Met	Phe	
					185					190					195
Gln	Leu	His	Gly	Asp	Glu	Phe	Cys	Pro	Val	Leu	Asp	Ala	Thr	Phe	
					200					205					210
Val	Met	Val	Ala	Arg	Asp	Ser	Glu	Asn	Lys	Gly	Pro	Ala	Phe	Val	
					215					220					225
Asn	Pro	Leu	Ile	Pro	Glu	Ser	Pro	Glu	Glu	Glu	Glu	Leu	Phe	Arg	
					230					235					240
Gln	Gly	Glu	Leu	Asn	Lys	Gly	Arg	Arg	Ile	Ala	Phe	Ser	Ser	Thr	
					245					250					255
Ser	Leu	Leu	Lys	Met	Ala	Pro	Ser	Ala	Glu	Glu	Arg	Thr	Thr	Ile	
					260					265					270
His	Glu	Met	Phe	Leu	Ser	Thr	Leu	Asp	Pro	Lys	Thr	Ile	Ser	Phe	
					275					280					285
Arg	Ser	Arg	Val	Leu	Pro	Ser	Asn	Ala	Val	Trp	Met	Glu	Asn	Ser	
					290					295					300
Lys	Leu	Lys	Ser	Leu	Glu	Ile	Cys	His	Pro	Gln	Glu	Arg	Asn	Ile	
					305					310					315
Phe	Asn	Arg	Ile	Phe	Gly	Gly	Phe	Leu	Met	Arg	Lys	Ala	Tyr	Glu	
					320					325					330
Leu	Ala	Trp	Ala	Thr	Ala	Cys	Ser	Phe	Gly	Gly	Ser	Arg	Pro	Phe	
					335					340					345
Val	Val	Ala	Val	Asp	Asp	Ile	Met	Phe	Gln	Lys	Pro	Val	Glu	Val	
					350					355					360
Gly	Ser	Leu	Leu	Phe	Leu	Ser	Ser	Gln	Val	Cys	Phe	Thr	Gln	Asn	
					365					370					

```
<210> 17
<211> 723
<212> DNA
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<223> Incyte ID No: 2293764CB1
```

<400> 17							
gcagcaacag	agttgcaggt	gtaaaataac	gggaaggcgg	gatgcgtggc	taaattgctc	60	
tgcgtgcaca	aagagtagga	gagcccagag	ttccagaatg	cccctaattc	cgaacaccac	120	
aggggtgagtc	tggagcaagt	cacctgggag	ggcttacagg	tgccataatg	aaggcctggg	180	
gcactgtggt	agtgacctg	gccacgctga	tgggtgtcac	tgtggatgcc	aagatctatg	240	
aactctgcga	gctggcgcca	agactggaga	gagcagggct	gaacggctac	aagggtctacg	300	
gcgttggaga	ctggctgtgc	atggctcatt	atgagagtgg	ctttgacacc	gccttcgtgg	360	
accacaatcc	tgatggcagc	agtgaatatg	gcattttcca	actgaattct	gcctgggtgg	420	
gtgacaatgg	cattacaccc	accaagaacc	tctgccacat	ggattgtcat	gacctgctca	480	
atccgcatat	tctggatgac	atcaggtgtg	ccaagcagat	tgtgtcctca	cagaatgggc	540	
ttcttgctg	gacttcttgg	aggtctacat	gttctggcca	tgatttatct	gaatggctca	600	
aggggtgtga	tatgcatgtg	aaaattgata	caaaaattca	tccatgactc	agattcgaag	660	
agacagattt	tatcttctct	tcatttcttc	atattgtcac	tttaataaag	gatgggtactc	720	
gtc						723	

```
<210> 18
<211> 1228
<212> DNA
<213> Homo sapiens
```

```
<220>
<221> misc_feature
<223> Incyte ID No: 949738CB1
```

<400> 18						
cccggagccg	ccagaccgtc	gcgccccctgc	cccatcgtag	tatatgagct	cgccacaca	60
aggacccccg	ctaaaagcca	gagctcccag	tccccgaggc	ttgaagacgg	ggactccctt	120
ctccaccaac	tctgtcctcg	gggggggtggg	gccccagccg	agatcacagc	gcgacaggag	180
tggggggtggc	cgctggagac	aggtgaagaa	acaagaaaac	taagaaatcc	gagcggttgg	240
aggggggagtc	tgtgtgggat	ggatggggac	gccggggggag	gggctgggcc	gctgctccca	300
tgccctgac	cggggagctc	cacagagcct	ggcgtcgggg	gaaggtgcgg	gggctggcct	360
tcccgtctcg	gactctgcca	aagetcaaa	ggagcacggg	gtgtctgggag	gtaaaactgag	420
gcaacgactg	gggctacagc	tgatcgaact	gccactctgag	gagtcattgc	cgtctgggacc	480
gctgcttggc	gacacggccg	tgatccaagg	ggacacggcc	ctaatacagc	ggccctggag	540
ccccgctcgt	aggccagagg	tcgatggagt	ccgcaaagcc	ctgcaagacc	tggggctccg	600
aatttgtggaa	ataggagacg	agaacgcgac	gctggatggc	actgacgttc	tcttcaccgg	660
ccgggagttt	ttcgtaggcc	tctccaaatg	gaccaatac	cgaggagctg	agatcgtggc	720
ggacacgttc	cgggacttcg	ccgtctccac	tgtgccagtc	tcgggtccct	cccacctgcg	780
cggctctctgc	ggcatggggg	gacctcgcac	tgttgtggca	ggcagcagcg	acgctgcccc	840
aaaggctgtc	cgggcaatgg	cagtgtctgac	agatacccca	tatgcctccc	tgacctccc	900
agatgacgca	cgtctgact	gtctcttct	tcgtcctggg	ttggctgggt	tgcccccttt	960
ctcctgcac	ctgtggagtg	gggatctgcg	caacagccag	gaggcactgc	agaagctctc	1020
tgatgtcacc	ctggtaacctg	tgtcctgtct	agaactggag	aaggccggcg	ccgggctcag	1080
ctccctctgc	ttggtgctca	gcacacgccc	ccacagctga	gggcctggcc	ttgggggtact	1140

gctggccagg ggtaggatag tataggaagt agaaggggaa ggagggttag atagagaatg 1200
ctgaataggc agtagttggg agagaggg 1228

```
<210> 19
<211> 2155
<212> DNA
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<223> Incyte ID No: 1297034CB1
```

<400> 19						
cggtctcgagc	tctgtttctctg	ttctactgccc	ccaggagagccc	ggcggggtccg	ggactcccgt	60
ccgtgccggt	gcggggcgccg	gcattgtggct	gtgggaggac	cagggcgggcc	tcttgggccc	120
tttctccttc	ctgtctgctag	tgtctgtctct	ggtgacgcgg	agcccggtca	atgcttgcct	180
cctcaccggc	agcctcttctg	ttctactgcg	cgtcttcagc	tttgagccgg	tgcctctctg	240
cagggccctg	caggtgctca	agccccggga	ccgcatttct	gccatcgccc	accgtggcgg	300
cagccacgac	gcgcccgaga	acacgctggc	ggccattcgg	caggcagcta	agaatggagc	360
aacaggcgtg	gagttggaca	ttgagtttac	ttctgacggg	attcctgtct	taatgcacga	420
taacacagta	gataggacga	ctgatgggac	tgggcgattg	tgtgatttga	catttgaaca	480
aattaggaag	ctaagctctg	cagcaaacca	cagactcagg	aatgatttcc	ctgatgaaaa	540
gactcctacc	ctgaagggaag	ctgttgcaga	gtgcctaaac	cataacctca	caatcttctt	600
tgatgtcaaa	ggccattgcac	acaaggctac	tgaggctcta	agaaaaatgt	atatggaatt	660
tcttcaactg	tataataata	gtgtgggtctg	ttcttctctg	ccagaagtta	tctacaagat	720
gagacaaaac	gacgggatg	taataacagc	attaactcac	agaccttga	gcctaagcca	780
tacaggagat	gggaaaccac	gctatgatac	tttctggaaa	cattttatat	ttgttatgat	840
ggacattttg	ctcgatttga	gcattgcataa	tattcttgtg	tacctgtgtg	gaatttcagc	900
tttctctcatg	caaaaggatt	ttgtatcccc	ggcctacttg	aagaagtgg	cagctaaagg	960
aatccaggtt	gttgggttga	ctgttaatac	ctttgatgaa	aagagttact	acgaatccca	1020
tcttggttcc	agctatatca	ctgacagcat	ggtagaagac	tgcgaacctc	acttctagac	1080
tttcacgggtg	ggacgaaacg	ggttcagaaa	ctgccagggg	cctcatacag	ggatatcaaa	1140
ataccctttg	tgttagccca	ggccctgggg	aatcagggtg	ctcacacaaa	tgcaatagtt	1200
ggtcactgca	tttttacctg	aacaaaagct	aaaccgggtg	ttgccaccat	gcaccatggc	1260
atgccagagt	tcaacactgt	tgtcttggaa	aactctgggtc	tgaaaaaacg	cacaagagcc	1320
cctgcctctgc	cctagctgag	gcacacaggg	agaccagctg	aggataagca	cagattgaat	1380
tgtacaattt	gcagatgcag	atgtaaatgc	atgggacatg	catgataact	cagagttgac	1440
attttaaaac	ttgccacact	tatttcaaat	atttgtactc	agctatgtta	acatgtactg	1500
tagacatcaa	acttgtggcc	atactaataa	aattatttaa	aggagcacta	aaggaaaact	1560
gtgtgcccaag	catcatatcc	taaggcatac	ggaatttggg	gaagccacca	tgcaatccag	1620
tgaggcttca	gtgtacagca	acaaaaatgg	tagggagggtc	ttgaagccaa	tgagggattt	1680
atagcatctt	gaatagagag	ctgcaaacca	ccagggggca	gagttgcact	ttccagggt	1740
ttttaggaag	ctctgcaaca	gatgtgatct	gatcataggc	aattagaact	ggaagaaact	1800
tccaaaaata	tctaggtttg	tcttcatttt	acaaatgagg	aaactaaaact	ctgtggaagg	1860
gaagggggtg	cctcaaaagt	cacagcttag	ctgggcacag	tggctcatgc	cgataatccc	1920
agcaatttcag	aaagctgagg	caggaggatt	acttgaggcc	agactggggc	atatagcaag	1980
accccatctc	taaaaaatta	ggcatgtgtg	tgcattgcctg	tattcccagc	tactcaggag	2040
gttgaggttg	gaggatcact	tgagcccaga	agttcaaggc	tgcaatgagc	catgattaca	2100
ccacggcact	acaaccttgg	tggcacagtg	agaacgcgac	tcttaaaaaa	aaaaa	2155

<210> 20
<211> 491
<212> DNA

<213> Homo sapiens

<220>

<221> misc feature

<223> Incyte ID No: 1553276CB1

<400> 20

gcccatggcc	gcagccctgg	cgctcgtggc	gggggtcctg	tggggggcgg	tgctgccct	60
ctggagcgcg	cttcgcaat	ataaaaagaa	aatcacagac	aggtgcttcc	accactctga	120
gtgctacagt	ggctgctgcc	tcatggactt	ggactccggt	ggagccttct	gtgccccag	180
ggccagaata	accatgatct	gcttgcccca	gtggttggaa	ctcttcaagg	gcagggattg	240
catcatatct	atctatgaag	cacctacccc	cagcttagta	tctgcacata	accaagggag	300
ctaccaacat	catctgccct	tgccggatgg	gcttgacgtg	catatccaag	gacttgatgt	360
gttcgcgcg	gtgccatatg	atttagagga	agatgcaggc	tggtcactgc	tcccttgggg	420
ccatagcccc	tggttgccac	caacttgctc	caaatccagc	tctgagaca	ttaaagtcac	480
ttcctgtcaa	a					497

<210> 21

$\langle 211 \rangle$ 2101

<212> DNA

<213> Homo sapiens

<220>

<221> misc feature

<223> Incyte ID No: 1702211CB1

<400> 21

ccacgcgctc	cgcttctgtc	gaaccagttg	taaggagaat	ggagagagca	gtgagagtgg	60
agtcgggggt	cctggtcggg	gtggtctgtc	tgtctctggc	atgccctgcc	acagccactg	120
ggcccgaagt	tgctcagcct	gaagtagaca	ccaccctggg	tctgtgtcga	ggccggcagg	180
tgggcgtgaa	gggcacagac	cgccttgtga	atgtctttct	gggcattcca	tttgcccagc	240
cgccactggg	ccctgaccgg	ttctcagccc	cacaccagc	acagccctgg	gaggggtgtg	300
gggatgccag	cactgcgcgc	ccaatgtgcc	tacaagacgt	ggagagcatg	aacagcagca	360
gatttgtcct	caacggaaaa	cagcagatct	tctccgtttc	agaggactcg	ctgggtcccca	420
acgtctatag	cccagctgag	gtccccgag	ggtccggtag	gccggtcatg	gtatgggtcc	480
atggaggcgc	tctgataact	ggcgtgtcca	cctctacga	tggatcagct	ctgggtgcct	540
atggggatgt	ggtcgtggtt	acagtccagt	accgccttgg	ggtccttggc	ttcttcagca	600
ctgggatgga	gcatgcacct	ggcaaccagg	gcttccctaga	tgtggtagct	gctttgcgct	660
gggtgcaaga	aaacatcgcc	cccttcgggg	gtgacctcaa	ctgtgtcact	gtctttggtg	720
gatctgccgg	tgggagcacc	atctctggcc	tggtcctgtc	cccagtggct	gcagggctgt	780
tccacagagc	catcacacag	agtggggcca	tcaccacccc	agggatcacc	gactctcacc	840
cttggccctc	agctcagaaa	atcggaata	ccttggcctg	cagctcccagc	tccccggctg	900
agatgggtgca	gtgccttcag	cagaaagaag	gagaagagct	ggtccttagc	aagaagctga	960
aaaatactat	ctatcctctc	accgttgatg	gcactgtctt	ccccaaaagc	ccaaggaac	1020
tctgaaggga	gaagcccttc	cactctgtgc	ccttccctcat	gggtgtcaac	aacctagagt	1080
tcagctggct	catccccagg	ggctggggtc	tcttggtatc	aatggagcag	atgagccggg	1140
aggacatgct	ggccatctca	acaccgctct	tgaccagtct	ggatgtgccc	cctgagatga	1200
tgcccaccgt	catagatgaa	tacctaggaa	gcaactcgga	cgcacaagcc	aaatgccagg	1260
cgttccagga	attcatgggt	gacgtattca	tcaatgttcc	caccgtcagt	ttttcaagat	1320
accttcgaga	ttctggaage	cctgtctttt	tctatgagtt	ccagatccga	cccagttctt	1380
ttgcgaagat	caaacctgcc	tgggtgaagg	ctgatcatgg	ggccgagggt	gctttttgtgt	1440
tcggaagtc	cttctctatg	gacgagagct	cccgcttggc	cttccagag	gccacagagg	1500
aggagaagca	gctaagcctc	accatgatgg	cccagtgga	ccactttgcc	ggcacagggg	1560
accccaatag	caaggctctg	cctccttggc	cccaattcaa	ccaggcggaa	caatatctgg	1620

21/30

WO 00/28045

PCT/US99/27009

<211> 753

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2011071CB1

<400> 23

```
atgcaggacg ctccccctgag ctgcctgtca ccgactaggt ggagcagtgt ttcttccgca 60
gactcaactg agaagtcagc ctctggggca ggcaccagga atctgccttt tcagttctgt 120
ctccggcagg ctttgaggat gaaggctgcg ggcattctga ccctcattgg ctgcctggtc 180
acaggcgccg agtccaaaat ctacactcgt tgcaaaactgg caaaaatatt ctcgagggct 240
ggcctggaca attactgggg cttcagcctt ggaaactgga tctgcatggc atattatgag 300
agcggctaca acaccacagc cccgacggtc ctggatgacg gcagcatcga ctatggcatc 360
ttccagatca acacgttcgc gtggtgcaga cgcggaaagc tgaaggagaa caaccactgc 420
catgtcgctt gctcagcctt gatcactgat gacctcacag atgcaattat ctgtgccagg 480
aaaattgtta aagagacaca aggaatgaac tattggcaag gctggaagaa acattgtgag 540
ggcagagacc tgtccgagtg gaaaaaaggc tgtgagggtt cctaaactgg aactggaccc 600
aggatgcttt gcagcaacgc cctaggattt gcagtgaatg tccaaatgcc tgtgtcatct 660
tgtcccgctt cctcccaata ttccttctca aacttggaag gggaaaatta agctatactt 720
ttaagaaaat aaatatcttc atttaaattg caa 753
```

<210> 24

<211> 1395

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2186517CB1

<400> 24

```
gccccagca tggcttggca gggctggccc gcggcggtggc agtgggtcgc cggtgctggg 60
ctcctcctcg tcttgtctct cgtcctactt gtgagcccc gcggctgccg agcgcggcgg 120
ggcctccgcg gtctgtctcat ggcgcacagc cagcggctgc tcttccgaat cgggtacagc 180
ctgtacaccc gcacctggct cgggtacctc ttctaccgac agcagctgcg cagggtcgcg 240
aatcgctacc ctaaaggcca ctcgaaaacc cagaccgcc tcttcaatgg agtgaagggtg 300
cttcccaccc ctgtcctctc ggacaactac agctacctca tcatcgacac ccaggcccag 360
ctggctgtgg ctgtggaccc ttcagaccct cgggctgtgc aggttccat tgaaggagaa 420
ggggtcacct tggctgccc tctgtgtact cacaagcact gggaccacag tggaggggaa 480
cgtgacctca gccggcgcca cgggactgt cgggtgtacg ggagccctca ggacggcatc 540
ccctacctca cccatccctt gtgtcatcaa gatgtggtca gcgtgggacg gcttcagatc 600
cggccctggg ctacacctgg ccacacacaa ggccatctgg tctacctact ggatggggag 660
ccctacaagg gtccctcctg cctcttctca ggggacctgc tcttctctc tggtgtggg 720
cggacctttg agggcaatgc agagaccatg ctgagctcac tggacactgt gctggggcta 780
ggggatgaca ccttctgtg gcctggtcat gagtatgcag aggagaacct gggctttgca 840
ggtgtgggtg agcccagaaa cctggccccg gagaggaaga tgcagtgggt gcagcggcag 900
cggctggagc gcaagggcac gtgcccactt accctgggag aggagcgtc ctacaacccg 960
ttcctgagaa cccactgcct ggcgctacag gaggtctctg ggccggggcc gggccccact 1020
ggggatgatg actactcccg ggcccagctc ctggaagagc tccgccggct gaaggatatg 1080
cacaagagca agtgatgccc ccagcgcccc cagcccagcc cactccccgc atggggaggc 1140
cgccaccacc aacacctcat catccttctc atcgctaaca ccaccacctc catcggcacc 1200
caagcgggca tcatcccccc aactgctca ggggagggga gggatcaggc gatgagactg 1260
```


tgaggcctaaa	agaaggggggc	ctgttgagg	ctgggaaccc	cgcagcgcga	ggctgcctca	1320
tcaacggcaa	gaggaaagga	ggggtctcgg	gacatctcca	gacctacca	actgggaggg	1380
tccctctctc	cttc					1395

```
<210> 25
<211> 1413
<212> DNA
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<223> Incyte ID No: 2253585CB1
```

<400>	25						
gcgagccggc	caacagcttg	caagcatgct	ccgctggacc	cgagcctgga	ggctcccgcg	60	
tgagggactc	ggccccacg	gccctagctt	cgcgaggggtg	cctgtcgcac	ccagcagcag	120	
cagcggcggc	cgagggggcg	ccgagccgag	gccgcttccg	ctttcctaca	ggcttctgga	180	
cggggaggca	gccctcccgg	ccgtcgtctt	tttgcaagg	ctcttcggca	gcaaaactaa	240	
cttcaactcc	atcgccaaga	tcttggecca	gcagacaggc	cgtaggggtgc	tgacgggtgga	300	
tgctcgtaac	cacgggtgaca	gccccacag	cccagacatg	agctacgaga	tcatgagcca	360	
ggacctgcag	gaccttctgc	cccagctggg	cctgggtgcc	tgcgtcgtcg	ttggccacag	420	
catgggagga	aagacagcca	tgctgctggc	actacagagg	ccagagctgg	tggaaactgt	480	
cattgctgta	gatatcagcc	cagtggaaag	cacaggtgtc	tcccactttg	caacctacgt	540	
ggcagccatg	agggccatca	acatcgcaga	tgagctgccc	cgctcccgtg	cccgaaaaact	600	
ggcggatgaa	gagctcagtt	ctgtctcca	ggacatggcc	gtgcggcagc	acctgtctac	660	
taacctggta	gaggtagacg	ggcgcttcgt	gtggagggtg	aacttggatg	ccctgaccca	720	
gcacctagac	aagatcttgg	ctttcccaca	gaggcaggag	tctacctcg	ggccaacact	780	
ctttctcctt	ggtggaaact	cccagttcgt	gcaccccagc	caccaccctg	agattatgcg	840	
gctcttccct	cggggccaga	tgcacaggtt	gccgaacgct	ggccactgga	tccacgctga	900	
ccgcccacag	gacttcatag	ctgccatccg	aggcttccctg	gtctaagagt	tgctggcaag	960	
aagatggccg	ggcgtggtgg	ctcatgcctg	taattccagc	actttgggag	gctaaggcgg	1020	
gaggatgact	tgaggccagg	agttggagac	cagcctggcc	aacatgggtga	aaccctgtct	1080	
ctactaaaaa	tacaaaaatt	agcctggcgt	ggtggtgcac	acctgtaatc	ccagctactc	1140	
gggaggctga	ggcaggagaa	tcacttgaac	cctggaggca	gaggttgcaa	tgagccgaga	1200	
tcacaccact	acactccagc	ctgggcacaa	gagcaagact	ctgtctcaaa	aaaaaacaaa	1260	
acaaaaagga	ggcacaaaac	cccaggcttc	aagtctctgc	agcctgctcc	acatttgggc	1320	
acagaaggac	tcagacaggc	actgtgtggg	cacgaggttt	tacaggggtg	agtcagacct	1380	
caggctttaa	tgaataaagc	actcagctat	aaa			1413	

```
<210> 26
<211> 1868
<212> DNA
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<223> Incyte ID No: 2447520CB1
```

<400>	26						
ttggagcctg	gagactgcta	gctgcctggt	tctttaagaa	ccagccctgg	tccagcccat	60	
tccgcaggcc	agcaagcttc	tgaaaagcaa	acctaggaag	tagctttcca	acataaagtg	120	
gaggtttcaa	cacaggagac	tttaagcaag	ttccagtgtg	tctatatattg	gtctggctga	180	
tcggctggac	tctggccttc	cccgcctcag	ttagcagaca	gctctgccct	agtgggcgct	240	

tagcctgcga	cggcagcccg	agaggatgtc	taacaagctt	ctttctcccc	accccatte	300
agttgttctc	aggtctgaat	tcaaaatggc	ctcatctect	gctgtccttc	gagcgtcccg	360
gctgtaccaa	tggagcctga	agagttcggc	gcagttcctg	gggtctccac	agctgaggca	420
ggttggtcag	atcattaggg	ttcctgctcg	gatggcggcg	acgctgatcc	tggagcctgc	480
gggccgctgc	tgctgggacg	aaccggtgcg	aatcgccgtg	cgcggcctag	ccccggagca	540
gccggtcacg	ctgcgcgcgt	ccttcgcgca	cgagaagggc	gcgcttttcc	agggccacgc	600
gcgctaccgc	gccgacctgc	ttggcgagct	ggacctggag	cgcgcgcccg	cgctggggcg	660
cagcttcgcg	gggcttagtc	ccatggggct	gctctggggc	ttggagcccg	agaaaccttt	720
ggtgcggctg	gtgaagcgcg	acgtgcgaac	gcccttggcc	gtggagctgg	aggtgctgga	780
tggccacgac	cccgaccccg	ggcggtctgt	gtgccagacg	cggcacgagc	gctacttctt	840
cccgccccgg	gtgcggcgcg	agccggtgcg	cgtggggccg	gtgcgaggca	cgctcttctt	900
gccgccagaa	cctgggccct	ttcctgggat	tgtggacatg	ttcggaactg	gaggtggcct	960
gctggagtat	cgggctagtc	tgctggctgg	gaagggtttt	gctgtgatgg	ctctggctta	1020
ttataactat	gaagacctcc	ccaagacctt	ggagacgctc	catctggagt	actttgaaga	1080
agccatgaac	tacttgctca	gtcatcccga	ggtaaaagg	ccaggagtgg	ggctgcttgg	1140
aatttccaaa	gggggtgagc	cttgcttttc	catggcctct	ttcctgaagg	gcatcacggc	1200
tgtgtctgtc	ctcaacggct	ctgtggccaa	tgttggggga	accttacgct	acaaggcgca	1260
gacctgcgcc	atcttggggc	tcaacagaaa	tgcatacaag	gtgaccaaa	atggctatgc	1320
agacatttgt	gatgtcctga	acagcccttt	ggaaggacct	gaccagaaga	gcttcatttc	1380
tgtggaaagg	gcagagagca	ccttctctgt	cctggtaggt	caggatgacc	acaactggaa	1440
gagtgagttc	tatgctaatt	aggcctgtaa	acgcttgacg	gcccatggga	ggagaaagcc	1500
ccagatcate	tgttaccacg	agacagggca	ctatatagag	cctccttact	tccccctgtg	1560
tccgggcttc	ctgcatgcct	tgggtgggcag	tcctattatc	tggggagggg	agcccagggc	1620
tcatgccatg	gctcaggtgg	atgcttggaa	acaactccag	actttcttcc	acaaacactt	1680
gggtggccac	gaggggacaa	tcccatcaaa	agtgtaaatt	ttatttgatc	atgtggccctc	1740
ctgtttgcta	atctctcctg	gaaacatctg	ccacatttag	tgtgtgtatg	tgtattcatt	1800
cttttgTTTT	taataactaa	agTTTTTTTc	cctcattatt	aaaatgaatt	taccagtaaa	1860
aaaaaaaa						1868

```
<210> 27
<211> 688
<212> DNA
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<223> Incyte ID No: 2481345CB1
```

<400> 27							
cggcgttact	atcaagcaac	caaactgcaa	gctttggggag	ttgttcgctg	tccctgccct		60
gctctgctag	ggagagaacg	ccagagcgga	ggcggctggc	ccggcggcag	gctctcagaa		120
ccgctaccgg	cgatgctact	gctgtgggtg	tcggtggctg	cagccttggc	gctggcggta		180
ctggcccccg	gagcagggga	gcagaggcgg	agagcagcca	aagcgcccaa	tgtggtgctg		240
gtcgtgagcg	actccttcga	tggaagggtta	acatttcctc	caggaagtca	ggtagtgaag		300
cttcctttta	tcaactttat	gaagacacgt	gggacttcct	ttctgaatgc	ctacacaaac		360
tctccaattt	gttgcccatc	acgcgcagca	atgtggagat	gcctcttcac	tcacttaaca		420
gaatcttgga	ataattttaa	gggtctagat	ccaaattata	caacatggat	ggatgtcatg		480
gagaggcatg	gctaccgaac	acagaaaattt	gggaaaactgg	actatacttc	aggacatcac		540
tccattagta	atcgtgtgga	agcgtgacaa	gagatgttgc	tttcttactc	agacaagaag		600
gcaggcccat	gggttaattct	atccgttaaca	ggactaaagt	cagagtgatg	gaaagggatt		660
ggcagaatac	agacaaaagca	gtaaactg					688

<210> 28

WO 00/28045

PCT/US99/27009

<211> 1375

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2484020CB1

<400> 28

```
gcgggggtggc ggcgggccggg cccccacggc ggcgggccgga gcagcagcag cagcagcagg 60
agccgcctc tatgatgaag ttcaagccca accagacgcg gacctacgac cgcgagggct 120
tcaagaagcg ggcgggcgtgc ctgtgcttcc ggagcgagca ggaggacgag gtgctgctgg 180
tgagttagcag ccggtaccca gaccagtggg ttgtcccagg aggaggaatg gaacccgagg 240
aggaacctgg cggtgctgcc gtgaggggaag tttatgagga ggctggagtc aaaggaaaac 300
taggcagact tctgggcata tttgagaacc aagaccgaaa gcacagaaca tatgtttatg 360
ttctaacagt cactgaaata ttagaagatt gggaagattc tgttaatat ggaaggaaga 420
gagagtgggt caaagtagaa gatgctatca aagttctcca gtgtcataaa cctgtacatg 480
cagagtatct ggaaaagcta aagctgggtt gtcccccage caatggaaat tctacagtcc 540
cttcccttcc ggataataat gccttgtttg taaccgctgc acagacctct gggttgccat 600
ctagtgtaa atagagagaa ctgggtaggc ctctcccacc atgtgcagtc tcatggggag 660
aggcttcttt cgtttcctcg tcaaacatct gattgacgct tgcaaactgt ctgaatttgc 720
catgcaaggt tttcaaaca tttgcatggt tttcagatgc tttcaaact ttttttaaaa 780
aaatagtgtg aaatatttta ataagccaaa gccatgtgga atttttgttt agatgcctta 840
actgtgccac accccacaac cccctatatt attttggttg tctattttct acagcatatt 900
ttcagttttt tgtccatttg acatcagtc gtggtttatt ttgtcatcag attacttgtg 960
ggtataccta ccccaaaatt gttttctcat tcacagcatt agcatattca gcaaatccat 1020
ctgtgggtgg aattaaaaat attattggta ttaaagaaat ccattcacc caaaacttgt 1080
tttacaggat tacaatttta attcaaaatt tccagatttg ggctatttct gtatgatcca 1140
ataacttatt ttgtcacagg gcttaatttg ccatttttgg ggatttgtcg actcattttg 1200
tctgaatttt cacaactggg attatgtcac tagctacctg atatggctat ttcccttata 1260
actcaatagt accttaacac aaagtataac tctgtagagt tgggtgaatat tttagggaaa 1320
tattagcaaa atgcatgtag taaagacatc ttatgaaaac tgtaaaaaaaa aaaaa 1375
```

<210> 29

<211> 1390

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2862528CB1

<400> 29

```
catcgctcca cctctgccct ccccttttat ggcgcgggccc gggetcatto attccgcgcc 60
gggcctgcca gacacctgcg cccttctgca gccgcccggc gcatccgcgc ccgcagcccc 120
cagcatgtcg ggcccagacg tcgagacgcc gtccgccatc cagatctgcc ggatcatgcg 180
gccagatgat gccaacgtgg ccggcaatgt ccacgggggg accatcctga agatgatcga 240
ggaggcaggc gccatcatca gcaaccggca ttgcaacagc cagaacgggg agcgtgtgt 300
ggccgcccctg gctcgtgtcg agcgcaccga ctctctgtct cccatgtgca tcggtgaggt 360
ggcgcatgtc agcgcggaga tcacctacac ctccaagcac tctgtggagg tgcaggtcaa 420
cgtgatgtcc gaaaacatcc tcacaggtgc caaaaagctg accaataagg ccacctgtg 480
gtatgtgccc ctgtcgctga agaattgtga caaggtcctc gaggtgcctc ctgttgtgta 540
ttcccggcag gagcaggagg aggagggccg gaagcggtat gaagcccaga agctggagcg 600
catggagacc aagtggagga acggggacat cgtccagcca gtcctcaacc caggtgtgac 660
```

WO 00/28045

PCT/US99/27009

```

catgaagctc atggatgagg tcgccgggat cgtggctgca cggcactgca agaccaacat 720
cgtcacagct tccgtggacg ccattaattt tcatgacaag atcagaaaag gctgcgtcat 780
caccatctcg ggacgcatga ccttcacgag caataagtc atggagatcg aggtgttggt 840
ggacgccgac cctgttggtg acagctctca gaagcgctac cgggcccga gtgccttctt 900
cacctacgtg tcgctgagcc aggaaggcag gtcgctgect gtgccccagc tgggtgcccga 960
gaccgaggac gagaagaagc gctttgagga aggcaaagg cggtagctgc agatgaaggc 1020
gaacgacagg gccacgcgga gcctcagccc tagactccct cctcctgcca ctggtgcctc 1080
gagtagccat ggcaacgggc ccagtgtcca gtcacttaga agttcccccc ttggccaaaa 1140
acccaattca cattgagagc tgggtgtgtc tgaagttttc gtatcacagt gttaacctgt 1200
actctctcct gcaaacctac acaccaaagc tttatttata tcattccagt atcaatgcta 1260
cacagtgttg tcccagagcg cgggaggcgt tgggcagaaa cctcgggaa tgcttccgag 1320
cacgctgtag ggtatgggaa gaaccagca ccactaataa agctgctgct tggctggaaa 1380
aaaaaaaaa 1390

```

<210> 30

<211> 3038

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3200650CB1

<400> 30

```

gcgcggctga gtgcggactg gagtgggaac cggggtcccc gcgcttagag aacacgcgat 60
gaccacgtgg agcctccggc ggaggccggc ccgcacgctg ggactcctgc tgctggctcg 120
cttgggcttc ctggtgctcc gcaggctgga ctggagcacc ctggtccctc tgcggctccg 180
ccatcgacag ctggggctgc aggccaaagg ctggaacttc atgctggagg attccacctt 240
ctggatcttc gggggctcca tccactattt ccgtgtgccc agggagtact ggagggaccg 300
cctgctgaag atgaaggcct gtggcttgaa caccctcacc acctatgttc cgtggaacct 360
gcatgagcca gaaagaggca aatttgactt tctctgggaa acttgacct tgaaggcctt 420
cgtcctgatg gccgcagaga tcgggctgtg ggtgattctg cgtccaggcc cctacatctg 480
tagtgagatg gacctcgggg gcttgcccag ctggctactc caagaccctg gcatgaggct 540
gaggacaact tacaaggcct tcaccgaagg agtggacctt tattttgacc acctgatgtc 600
cagggtggtg ccactccagt acaagcgtgg gggacctatc attgccgtgc aggtggagaa 660
tgaatatggt tcctataata aagacccgcg atacatgccc tacgtcaaga aggcactgga 720
ggaccgtggc attgtggaac tgctcctgac ttcagacaac aaggatgggc tgagcaaggg 780
gattgtccag ggagtcttgg ccaccatcaa ctgacagtca acacacgagc tgcagctact 840
gaccaccttt ctcttcaacg tccaggggac tcagcccaag atggtgatgg agtactggac 900
ggggtggttt gactcgtggg gaggccctca caatatcttg gattcttctg aggttttgaa 960
aaccgtgtct gccattgtgg acgcggctc ctccatcaac ctctacatgt tccacggagg 1020
caccaacttt ggcttcatga atggagccat gcacttccat gactacaagt cagatgtcac 1080
cagctatgac tatgatgctg tgctgacaga agccggcgat tacacggcca agtacatgaa 1140
gcttcgagac ttcttcggct ccattcagg catcctctc cctccccac ctgaccttct 1200
tcccagatg ccgtatgagc ccttaacgac agtcttgtag ctgtctctgt gggacgcctt 1260
caagtacctg ggggagccaa tcaagtctga aaagcccatc aacatggaga acctgccagt 1320
caatggggga aatggacagt ccttcgggta cattctctat gagaccagca tcacctcgtc 1380
tggcatcctc agtggccacg tgcatgatcg ggggcagggt tttgtgaaca cagtatccat 1440
aggattcttg gactacaaga caacgaagat tgctgtcccc ctgatccagg gttacaccgt 1500
gctgaggatc ttggtggaga atcgtgggag agtcaactat ggggagaata ttgatgacca 1560
gcgcaaaggc ttaattggaa atctctatct gaatgattca cccctgaaaa acttcagaat 1620
ctatagcctg gatatgaaga agagcttctt tcagagggtc ggcctggaca aatggagttc 1680
cctcccagaa acaccacat tacctgcttt cttcttgggt agcttgtcca tcagctccac 1740
ccctgtgac acctttctga agctggaggg ctgggagaa ggggttgtat tcacatgag 1800

```

WO 00/28045

PCT/US99/27009

```

ccagaacctt ggacgttact ggaacattgg accccagaag acgctttacc tcccagggtcc 1860
ctggttgagc agcggaatca accagggtcat cgtttttgag gagacgatgg cgggccctgc 1920
attacagttc acggaaaccc cccacctggg caggaaccag tacattaagt gagcgggtggc 1980
acccccctct gctggtgccca gtgggagact gccgcctcct cttgacctga agcctgggtgg 2040
ctgctgcccc accctcact gcaaaagcat ctcttaagt agcaacctca gggactgggg 2100
gctacagtct gccctgtct cagctcaaaa ccctaagcct gcaggggaaag gtgggatggc 2160
tctgggcctg gctttgttga tgatggcttt cctacagccc tgcctctgtg ccgaggctgt 2220
cgggctgtct ctaggggtgg agcagctaata cagatcgccc agcctttggc cctcagaaaa 2280
agtgtgaaa cgtgcccttg caccggacgt cacagccctg cgagcatctg ctggactcag 2340
gcgtgctctt tgctggttcc tgggaggctt ggccacatcc ctcatggccc cattttatcc 2400
ccgaaatcct ggggtgtgca ccagtgtaga ggggtgggaa ggggtgtctc acctgagctg 2460
actttgttct tccttcacaa cctcttgagc cttctttggg attctggaag gaactcggcg 2520
tgagaaacat tggacttccc ctttcccttc ccactcgtg cttcccacag ggtgacaggc 2580
tgggctggag aaacagaaat cctcaccctg cgtcttccca agtttagcagg tgtctctggc 2640
gttcagttag gaggacatgt gagtccctggc agaagccatg gcccatgtct gcacatccag 2700
ggaggaggac agaaggccca gctcacatgt gagtccctggc agaagccatg gcccatgtct 2760
gcacatccag ggaggaggac agaaggccca gctcacatgt gagtccctggc agaagccatg 2820
gcccatgtct gcacatccag ggaggaggac agaaggccca gctcagtgcc ccccgccccc 2880
cacccccac gcccgaaacag caggggcaga gcagccctcc ttcgaagtgt gtccaagtcc 2940
gcatttgagc cttgttctgg gggccagccc aacacctggc ttgggctcac tgtcctgagt 3000
tgcagtaaaag ctataacctt gaatcacaaa aaaaaaaa 3038

```

<210> 31

<211> 1340

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4107621CB1

<400> 31

```

gcgacctagc caggcgtgag ggagtgtacag cagcgcattc gcgggacgag agcgatgagt 60
gagaacgccg caccagggtct gatctcagag ctgaagctgg ctgtgccctg gggccacatc 120
gcagccaaag cctggggctc cctgcagggc cctccagttc tctgcctgca cggctggctg 180
gacaatgcca gctccttcga cagactcatc cctcttctcc cgcaagactt ttattacgtt 240
gccatggatt tcggagggtca tgggctctcg tcccattaca gcccagggtg cccatattac 300
ctccagactt ttgtgagtga gatccgaaga gttgtggcag ccttgaaatg gaatcgattc 360
tccattctgg gccacagctt cgggtggcgtc gtgggcggaa tgtttttctg taccttcccc 420
gagatgggtg ataaacttat cttgctggac acgccgctct ttctcctgga atcagatgaa 480
atggagaact tgctgacctc caagcggaga gccatagagc acgtgctgca ggtagaggcc 540
tcccaggagc cctcgacagt gttcagcctg aagcagctgc tgcagagggt actgaagagc 600
aatagccact tgagttagga gtgcggggag cttctcctgc aaagaggaa cagcaagggt 660
gccacaggte tggttctgaa cagagaccag aggtcgcctt gggcagagaa cagcattgac 720
ttcatcagca gggagctgtg tgcgcatctc atcaggaagc tgcaggccca tgtcctgttg 780
atcaaagcag tccacggata ttttgattca agacagaatt actctgagaa ggagtccctg 840
tcgttcatga tagacacgat gaaatccacc ctcaaagagc agttccagtt tgtggaagtc 900
ccaggcaatc actgtgtcca catgagcgaa cccagcacg tggccagtat catcagctcc 960
ttcttacagt gcacacacat gctcccagcc cagctgtagc tctgggcctg gaactatgaa 1020
gacctagtgc tcccagactc aacactggga ctctgagttc ctgagcccca caacaaggcc 1080
agggatgggt gggacaggcc tcactagtct tgaggccag cctaggatgg tagtcagggg 1140
aaggagcgag attccaactt caacatctgt gacctcaaga gggagacaga gtctgggttc 1200
cagggtgct ttctcctggc taataataaa tatccagcca gctggaggaa ggaagggcag 1260
gctgggcccc ctagcctttt cctgctgccc caactggatg gaaaataaaa ggttcttgta 1320

```

WO 00/28045

PCT/US99/27009

ttctcaaaaa aaaaaaaaaa

1340

<210> 32

<211> 1717

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4661133CB1

<400> 32

```

cgcccccgga caccgctgtc cgcctcccgg gctgtcctca gcaagggcgc ggtctggtac 60
tcgtgcgtct tttatgcct cagtttccct ccgcgcgacta gcgcgcgggg cccggttctc 120
catcgcgcg acggcagcct agcgcaatga ggcgggcagc actgcggctt tgtgccttgg 180
gcaaagggca gcttactcct ggaagaggac tgactcaagg accccagaac cccaagaaac 240
agggaatctt ccacattcat gaagcatgtt catctataca tgtgaatcat gttcgagata 300
agttgcgggg gatagtagga gcatccacaa actggagaga ccatgtgaag gcaatggaag 360
aaaggaaatt acttcatagt ttcttggtta aatcacagga tggactgcct cctaggagaa 420
tgaaggacag ttatatgaa gttctcttgc ctttgggcag tgagcctgaa ttacgagaga 480
aatatttgac tgttcaaaac accgtaagat ttggcaggat tcttgaggat cttgacagct 540
tgggagttct tatttggtac atgcacaaca aaatccactc cgccaagatg tctcctttat 600
cgatagttac agccctggtg gataagattg atatgtgtaa gaagagcttg agcccagaac 660
aggacattaa gttcagtggt catgttagct gggtcgggaa gacatccatg gaagtgaaga 720
tgcaaattgt ccagttacat ggtgatgaat tttgtcctgt tttggatgca acatttgtaa 780
tggtggctcg tgattctgaa aataaagggc cggcatttgt aaatccactc atccctgaaa 840
gccagagga agaggagctc tttagacaag ggaattgaa caaggggaga agaattgcct 900
tcagctccac gtcgttactg aaaatggccc ccagcgtga ggagaggacc accatacatg 960
agatgtttct cagcacactg gatccaaaga ctataagttt tcggagtcga gttttaccct 1020
ctaattgcagt gtggatggag aattcaaaac tgaagagttt ggaaatttgc caccctcagg 1080
agcggaaacat tttcaatcgg atctttggtg gtttccttat gaggaaggca tatgaacttg 1140
cgtgggctac tgcttgtagc tttggtggtt ctgcaccgtt tgtggtagca gtagatgaca 1200
tcatgtttca gaaacctgtt gaggttggct cattgctctt tctttcttca caggatgct 1260
ttactcagaa taattatatt caagtcagag tacacagtga agtggcctcc ctgcaggaga 1320
agcagcatat aaccaccaat gtctttcatt tcacgttcat gtcggaaaaa gaagtgccat 1380
tggttttccc aaaaacatat ggagagtcca tggtgtactt agatgggcag cggcatttca 1440
actccatgag tggcccagcg accttgagaa aggactacct tgtggagccc taagaacacc 1500
acatttggtg aaaactagca ctctaccac agtgacgtgg tatctgatga agacctgatc 1560
gagtgatttg attttagtat tgcttcgtgt cctccacaca ggaggaggat gtattcagcc 1620
tttaggatga tcagaaaagc agaaagagag agtggccgga tggggctgag gggagaaaga 1680
attattaaac aataaatact ttcaagacaa aaaaaaa 1717

```

<210> 33

<211> 148

<212> PRT

<213> Colobus guereza

<300>

<308> GenBank ID No: g1790927

<400> 33

```

Met Lys Ala Leu Ile Ile Leu Gly Leu Val Leu Leu Ser Val Thr
  1             5             10            15
Val Gln Gly Lys Ile Phe Glu Arg Cys Glu Leu Ala Arg Thr Leu

```

				20					25					30
Lys	Lys	Leu	Gly	Leu	Asp	Gly	Tyr	Lys	Gly	Val	Ser	Leu	Ala	Asn
				35					40					45
Trp	Val	Cys	Leu	Ala	Lys	Trp	Glu	Ser	Gly	Tyr	Asn	Thr	Asp	Ala
				50					55					60
Thr	Asn	Tyr	Asn	Pro	Gly	Asp	Glu	Ser	Thr	Asp	Tyr	Gly	Ile	Phe
				65					70					75
Gln	Ile	Asn	Ser	Arg	Tyr	Trp	Cys	Asn	Asn	Gly	Lys	Thr	Pro	Gly
				80					85					90
Ala	Val	Asn	Ala	Cys	His	Ile	Ser	Cys	Asn	Ala	Leu	Leu	Gln	Asn
				95					100					105
Asn	Ile	Ala	Asp	Ala	Val	Ala	Cys	Ala	Lys	Arg	Val	Val	Ser	Asp
				110					115					120
Pro	Gln	Gly	Ile	Arg	Ala	Trp	Val	Ala	Trp	Lys	Lys	His	Cys	Gln
				125					130					135
Asn	Arg	Asp	Val	Ser	Gln	Tyr	Val	Glu	Gly	Cys	Gly	Val		
				140					145					

```
<210> 34
<211> 148
<212> PRT
<213> Colobus angolensis
```

<300>
<308> GenBank ID No: q1790967

<400>	34														
Met	Lys	Ala	Leu	Ile	Ile	Leu	Gly	Leu	Val	Leu	Leu	Ser	Val	Thr	
1				5					10					15	
Val	Gln	Gly	Lys	Ile	Phe	Glu	Arg	Cys	Glu	Leu	Ala	Arg	Thr	Leu	
				20					25					30	
Lys	Lys	Leu	Gly	Leu	Asp	Gly	Tyr	Lys	Gly	Val	Ser	Leu	Ala	Asn	
				35					40					45	
Trp	Val	Cys	Leu	Ala	Lys	Trp	Glu	Ser	Gly	Tyr	Asn	Thr	Asp	Ala	
				50					55					60	
Thr	Asn	Tyr	Asn	Pro	Gly	Asp	Glu	Ser	Thr	Asp	Tyr	Gly	Ile	Phe	
				65					70					75	
Gln	Ile	Asn	Ser	Arg	Tyr	Trp	Cys	Asn	Asn	Gly	Lys	Thr	Pro	Gly	
				80					85					90	
Ala	Val	Asn	Ala	Cys	His	Ile	Ser	Cys	Asn	Ala	Leu	Leu	Gln	Asn	
				95					100					105	
Asn	Ile	Ala	Asp	Ala	Val	Ala	Cys	Ala	Lys	Arg	Val	Val	Ser	Asp	
				110					115					120	
Pro	Gln	Gly	Ile	Arg	Ala	Trp	Val	Ala	Trp	Lys	Lys	His	Cys	Gln	
				125					130					135	
Asn	Arg	Asp	Val	Ser	Gln	Tyr	Val	Glu	Gly	Cys	Gly	Val			
				140					145						

```
<210> 35
<211> 148
<212> PRT
<213> Nasalis larvatus
```

```
<300> misc_feature
```

<308> GenBank ID No: g1790984

<400> 35

Met	Lys	Ala	Leu	Ile	Ile	Leu	Gly	Leu	Val	Leu	Leu	Ser	Val	Thr	15
1				5					10						
Val	Gln	Gly	Lys	Ile	Phe	Glu	Arg	Cys	Glu	Leu	Ala	Arg	Thr	Leu	30
				20					25						
Lys	Lys	Leu	Gly	Leu	Asp	Gly	Tyr	Lys	Gly	Val	Ser	Leu	Ala	Asn	45
				35					40						
Trp	Val	Cys	Leu	Ala	Lys	Trp	Glu	Ser	Gly	Tyr	Asn	Thr	Glu	Ala	60
				50					55						
Thr	Asn	Tyr	Asn	Pro	Gly	Asp	Glu	Ser	Thr	Asp	Tyr	Gly	Ile	Phe	75
				65					70						
Gln	Ile	Asn	Ser	Arg	Tyr	Trp	Cys	Asn	Asn	Gly	Lys	Thr	Pro	Gly	90
				80					85						
Ala	Val	Asp	Ala	Cys	His	Ile	Ser	Cys	Ser	Ala	Leu	Leu	Gln	Asn	105
				95					100						
Asn	Ile	Ala	Asp	Ala	Val	Ala	Cys	Ala	Lys	Arg	Val	Val	Ser	Asp	120
				110					115						
Pro	Gln	Gly	Ile	Arg	Ala	Trp	Val	Ala	Trp	Arg	Asn	His	Cys	Gln	135
				125					130						
Asn	Arg	Asp	Val	Ser	Gln	Tyr	Val	Lys	Gly	Cys	Gly	Val			
				140					145						